AMERICAN JOURNAL OF MATHEMATICS

FOUNDED BY THE JOHNS HOPKINS UNIVERSITY

EDITED BY

E. W. CHITTENDEN
UNIVERSITY OF IOWA

A. B. COBLE
UNIVERSITY OF ILLINOIS

ABRAHAM COHEN
THE JOHNS HOPKINS UNIVERSITY

G. C. EVANS
RICE INSTITUTE

F. D. MURNAGHAN
THE JOHNS HOPKINS UNIVERSITY

WITH THE COÖPERATION OF

FRANK MORLEY
E. T. BELL
W. A. MANNING
HARRY BATEMAN
J. R. KLINE
E. P. LANE
HARRY LEVY
MARSTON MORSE

PUBLISHED UNDER THE JOINT AUSPICES OF

THE JOHNS HOPKINS UNIVERSITY
AND
THE AMERICAN MATHEMATICAL SOCIETY

Volume LII, Number 4

OCTOBER, 1930

THE JOHNS HOPKINS PRESS
BALTIMORE, MARYLAND
U. S. A.

CONTENTS

THE AMERICAN JOURNAL OF MATHEMATICS will appear four times yearly.

The subscription price of the JOURNAL for the current volume is \$7.50 (foreign postage 25 cents); single numbers \$2.00.

A few complete sets of the JOURNAL remain on sale.

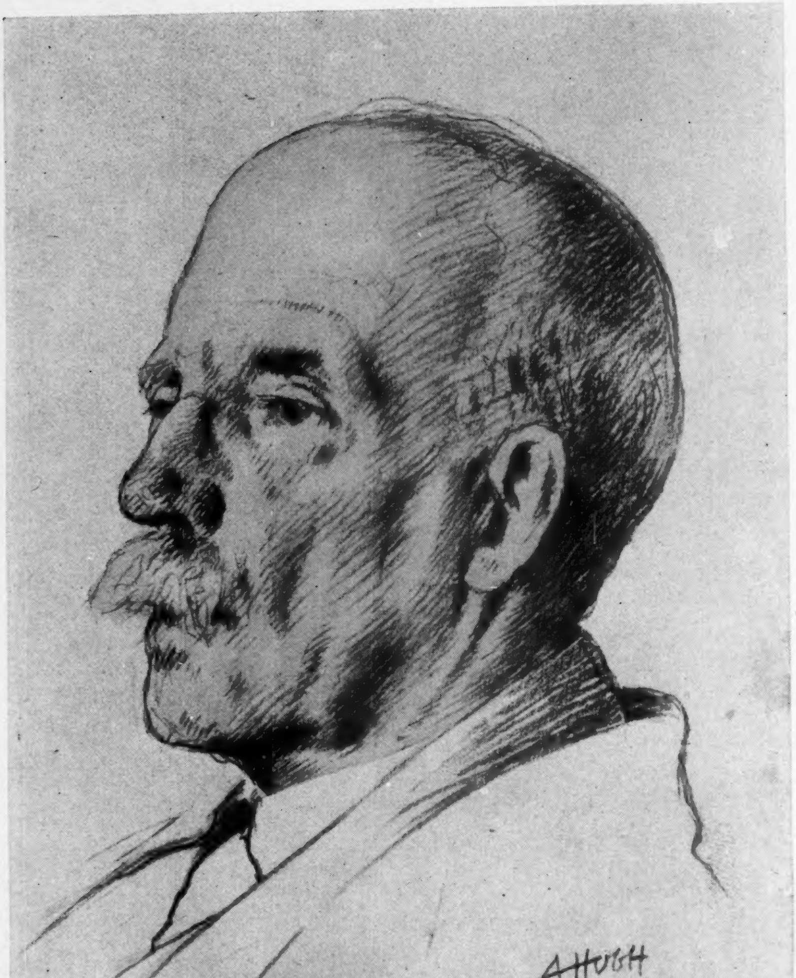
Papers intended for publication in the JOURNAL may be sent to any of the Editors.

Editorial communications may be sent to Dr. A. COHEN at The Johns Hopkins University.

Subscriptions to the JOURNAL and all business communications should be sent to THE JOHNS HOPKINS PRESS, BALTIMORE, MARYLAND, U. S. A.

Entered as second-class matter at the Baltimore, Maryland, Postoffice, acceptance for mailing at special rate of postage provided for in Section 1103, Act of October 3, 1917, Authorized on July 3, 1918

PRINTED BY THE J. H. FURST COMPANY
BALTIMORE, MD.



Frank Morley

The Problem of Lagrange in the Calculus of Variations.

By Gilbert Ames Bliss.

TABLE OF CONTENTS.

INTRODUCTION.

The problem of the calculus of variations principally considered in this paper is that of finding in a class of arcs

$$(1)\qquad y_i = y_i(x) \qquad (x_1 \leqq x \leqq x_2;\ i = 1, \cdots, n)$$

satisfying a set of differential equations

$$(2)\qquad \phi_\alpha(x, y_1, \cdots, y_n, y_1', \cdots, y_n') = 0 \qquad (\alpha = 1, \cdots, m < n)$$

and joining two fixed points in the space of points $(x, y_1, \cdots, y_n)$, one which minimizes an integral of the form

$$(3)\qquad I = \int_{x_1}^{x_2} f(x, y_1, \cdots, y_n, y_1', \cdots, y_n')\, dx.$$

A number of paragraphs are also devoted to the similar problem for which the end-points are variable.

The problem seems to have been first formulated by Lagrange for the general case here studied, though somewhat less precisely than in the statement above. He also gave the multiplier rule described in Section 5 below which had been previously deduced by Euler and himself for a number of more special cases. Important additions to the theory have been made by Clebsch, A. Mayer, Kneser, Hilbert, von Escherich, Hahn, Bolza, and many others. Comprehensive treatments of the problem have been given by Bolza [3] * and Hadamard [4], that of Bolza being the more complete. In Chapter V below a brief sketch of the history of the problem is given with a bibliography of the more important papers on which the text of this paper is based.

Since the literature of the problem is extensive and widely scattered, and since recent developments make possible important simplifications, even as compared with the excellent treatments of Bolza and Hadamard, it seemed justifiable to the author of this paper to attempt anew the presentation of those parts of the theory leading to the necessary conditions for a minimum, and to those sufficient to insure a minimum. The paper is a record of lectures which the author has given at intervals for some years past at the University of Chicago.

Some special features of the methods used may perhaps be mentioned. The deduction of the Euler-Lagrange multiplier rule in Sections 3-5 is based upon suggestions in papers by Hahn [13, p. 271] and the author [16, pp. 307, 312], but is different from the proofs hitherto given. The definition of

* The figures in the square brackets refer to the bibliographical list at the end of the paper.

normal arcs in Sections 7 and 8 is that of Bolza [19, p. 440]. A new application of the definition, in Section 15, makes it possible to deduce without the use of special methods the multiplier rule for the case when the functions ϕ_α contain none of the derivatives y_i', as a corollary to the rule deduced in Section 5. The discussions of the necessary conditions of Weierstrass and Clebsch, and of the envelope theorem with the associated deduction of the necessary condition of Mayer, are essentially those of Hahn [21] and Bolza [3, pp. 603-10], but are greatly simplified by the use of the auxiliary formulas of Section 21. The analytic proof of the necessary condition of Mayer in Section 26, by means of the minimum problem associated with the second variation, was suggested by the author for simpler cases [27] and applied to the problem of Lagrange by D. M. Smith [28]. By means of the theory of the minimum problem of the second variation the very elaborate theories of that variation due to Clebsch [29], von Escherich [31], Hahn [33], and others, can be much simplified, as the author has shown [35]. The applications important for this paper are in Sections 26 and 32. The theory of Mayer fields in Sections 28 and 29, and the proofs of the sufficiency theorems in Sections 30 and 31, have been simplified as far as seemed possible.

An effort has been made in each theorem to state clearly the underlying hypotheses. The proof of the multiplier rule in Section 5, for example, is independent of the assumption that the determinant R of page 11 is different from zero. In many of the succeeding theorems, however, this assumption is either made explicitly or else is a consequence of the property III′ which appears frequently.

CHAPTER I.

THE EULER-LAGRANGE MULTIPLIER RULE.

1. *Hypotheses.* In this first chapter the famous multiplier rule of Euler and Lagrange, describing the differential equations satisfied by a minimizing arc for the problem of Lagrange stated in the introduction, is to be deduced. For convenience in the following pages the set $(x, y_1, \cdots, y_n, y_1', \cdots, y_n')$ will be represented by (x, y, y').

As usual we concentrate attention on a particular arc E_{12} with the equations (1) and inquire what properties it must have if it is to be a minimizing arc. The analysis is based upon the following hypotheses:

(a) the functions $y_i(x)$ defining E_{12} are continuous on the interval x_1x_2 and this interval can be subdivided into a finite number of parts on each of which the functions have continuous derivatives;

(b) in a neighborhood $\mathfrak{N}$ of the values (x, y, y') on the arc E_{12} the functions f, ϕ_a have continuous derivatives up to and including those of the **fourth order;**

(c) at every element (x, y, y') on E_{12} the $m \times n$-dimensional matrix $\| \phi_{ay_i'} \|$ has rank m.

The subscript y_i' here indicates the partial derivative of ϕ_a with respect to y_i'. In the following pages literal subscripts, following the indices of functions and elsewhere, will be frequently used to indicate partial derivatives. The hypothesis (c) implies that the equations $\phi_a = 0$ are all independent near E_{12} when regarded as functions of the variables y_i'.

2. *Examples.* A common example of a Lagrange problem is that of the brachistochrone in a resisting medium [3, p. 5]. The differential equation of the motion [5, p. 44] becomes for this case

$$dv/dt = d^2s/dt^2 = g(dy/ds) - R(v),$$

where $R(v)$ is the retardation on the particle per unit mass due to the resistance of the medium. Multiplying by $ds/dx = (ds/dt)(dt/dx) = v dt/dx$ we find the equation

$$vv' = gy' - R(v)s' = gy' - R(v)(1 + y'^2)^{1/2} \tag{4}$$

where the primes denote derivatives with respect to x. The problem is then to find among the pairs of functions $y(x)$, $v(x)$ which have the end-values

$$y(x_1) = y_1, \quad v(x_1) = v_1, \quad y(x_2) = y_2$$

and satisfy equation (4) one which minimizes the time integral

$$I = \int_{s_1}^{s_2} (ds/v) = \int_{x_1}^{x_2} (1/v)(1 + y'^2)^{1/2} dx.$$

It should be noted that this problem is not precisely like that stated in section 1 since the value of v is not prescribed at $x = x_2$. It is in fact a problem of Lagrange with second end-point variable.

The so-called isoperimetric problems form a very large class, and all of them may be stated as Lagrange problems. For example we may seek to find among the arcs $y = y(x)$ $(x_1 \leqq x \leqq x_2)$, joining two given points and having a given length, one which has its center of gravity the lowest. This is the problem of determining the form of a hanging chain suspended between two pegs at its ends. Analytically the problem is to find among the functions $y(x)$ $(x_1 \leqq x \leqq x_2)$ satisfying the conditions

$$y(x_1)=y_1,\quad y(x_2)=y_2,\quad \int_{x_1}^{x_2}(1+y'^2)^{1/2}dx=l$$

one which minimizes the integral

$$I=\int_{x_1}^{x_2}y(1+y'^2)^{1/2}dx. \tag{5}$$

This problem may be made over into one of the Lagrange type by introducing the new variable

$$z(x)=\int_{x_1}^{x}(1+y'^2)^{1/2}dx$$

satisfying the differential equation $z'=(1+y'^2)^{1/2}$. The problem is then to find among the pairs $y(x)$, $z(x)$ satisfying $y(x_1)=y_1$, $z(x_1)=0$, $y(x_2)=y_2$, $z(x_2)=l$, $z'=(1+y'^2)^{1/2}$ one which minimizes the integral (5).

More generally suppose we wish to find among the functions $y(x)$ satisfying

$$y(x_1)=y_1,\quad y(x_2)=y_2$$

$$\int_{x_1}^{x_2}g(x,y,y')dx=k,\quad \int_{x_1}^{x_2}h(x,y,y')dx=l$$

one which minimizes

$$I=\int_{x_1}^{x_2}f(x,y,y')dx. \tag{6}$$

The problem is equivalent to that of finding among the sets of functions $y(x)$, $u(x)$, $v(x)$ satisfying

$$\begin{aligned}
&y(x_1)=y_1, && u(x_1)=0, && v(x_1)=0,\\
&y(x_2)=y_2, && u(x_2)=k, && v(x_2)=l,\\
&u'=g(x,y,y'), && && v'=h(x,y,y'),
\end{aligned}$$

one which minimizes the integral (6). Evidently a similar transformation of the problem could be made no matter how many isoperimetric integrals were to have prescribed constant values.

These illustrations suffice to show the wide applicability of the Lagrange problem.

3. *Admissible arcs and variations.* An *admissible arc*

$$y_i=y_i(x)\qquad (i=1,\cdots,n;\ x_1\leqq x\leqq x_2) \tag{7}$$

is one with the continuity properties (a) of Section 1, whose elements (x,y,y') all lie in the region $\mathfrak{R}$, and which satisfies the equations $\phi_\alpha=0$. If a one-parameter family of admissible arcs

$$y_i = y_i(x, b) \qquad (i = 1, \cdots, n) \tag{8}$$

containing a particular admissible arc E_{12} for the parameter value $b = b_0$ is given, the functions

$$\eta_i(x) = y_{ib}(x, b_0) \qquad (i = 1, \cdots, n),$$

where the subscript b indicates as usual a partial derivative of $y_i(x, b)$, are called variations of the family along E_{12}.

In the tensor analysis it is agreed that a product $G_{ik}H_k$ shall stand for the sum $\Sigma_k G_{ik}H_k$. In other words, when an index k occurs twice in the same term it is understood that the term really represents the sum of n terms of the same type. The index with respect to which the sum is taken is called an umbral index.

With this convention in mind we may define for the arc E_{12} mentioned above the so-called *equations of variation* by the formula

$$\Phi_\alpha(x, \eta, \eta') = \phi_{\alpha y_i}\eta_i + \phi_{\alpha y_i'}\eta_i' = 0 \qquad (\alpha = 1, \cdots, m) \tag{9}$$

in which i is an umbral index with the range $1, \cdots, n$, and the coefficients $\phi_{\alpha y_i}$, $\phi_{\alpha y_i'}$ are supposed to have as arguments the functions $y_i(x)$ belonging to E_{12}. These equations are satisfied by the variations $\eta_i(x)$ along E_{12} as we may readily see by substituting the functions (8) in the equations $\phi_\alpha = 0$, differentiating for b, and setting $b = b_0$. A set of functions $\eta_i(x)$ with the continuity properties described in (a) of Section 1 and satisfying the equations of variation (9) is called a *set of admissible variations,* a nomenclature which is justified by the following very important theorem:

Satisfied all b.

For every set of admissible variations $\eta_i(x)$ along the admissible arc E_{12} there exists a one parameter family (8) of admissible arcs containing E_{12} for the value $b = 0$ and having the functions $\eta_i(x)$ as its variations along E_{12}. For this family the functions $y_i(x, b)$ are continuous and have continuous derivatives with respect to b for all values (x, b) near those defining E_{12}, and the derivatives $y_{ix}(x, b)$ have the same property except possibly at the values of x defining corners of E_{12}.

To prove this theorem we enlarge the system $\phi_\alpha = 0$ to have the form

$$\phi_1 = 0, \cdots, \phi_m = 0, \quad \phi_{m+1} = z_{m+1}, \cdots, \phi_n = z_n \tag{10}$$

where $z_{m+1}, \cdots, z_n$ are new variables and $\phi_{m+1}, \cdots, \phi_n$ are new functions of x, y, y' such that the functional determinant $|\partial\phi_i/\partial y_k'|$ is different from zero along E_{12}.* By means of the last $n - m$ of these equations the functions

* For a proof of the possibility of this adjunction see Bliss [16, pp. 307, 312].

$y_i(x)$ belonging to E_{12} define a set of functions $z_r(x)$ $(r = m + 1, \cdots, n)$. We have a corresponding system of equations of variation

$$(11) \qquad \Phi_1 = 0, \cdots, \Phi_m = 0, \quad \Phi_{m+1} = \zeta_{m+1}, \cdots, \Phi_n = \zeta_n$$

along E_{12}, the last $n - m$ of which define a set $\zeta_r(x)$ $(r = m + 1, \cdots, n)$ corresponding to every set of admissible variations $\eta_i(x)$.

Suppose now that the set $\eta_i(x)$ is an admissible set of variations for E_{12} defining a set $\zeta_r(x)$ by means of equations (11). Since the functional determinant $| \partial\phi_i/\partial y_k' |$ is different from zero along E_{12} the existence theorems for differential equations * tell us that the system

$$(12) \quad \phi_\alpha = 0, \quad \phi_r = z_r(x) + b\zeta_r(x) \quad (\alpha = 1, \cdots, m;\ r = m + 1, \cdots, n)$$

determines uniquely a one-parameter family of solutions $y_i = y_i(x, b)$ with the initial values $y_i(x_1) + b\eta_i(x_1)$ at $x = x_1$. This family contains E_{12} for $b = 0$ and has variations which have the initial values $\eta_i(x_1)$ at $x = x_1$ and which satisfy the equations (11) with the functions $\zeta_r(x)$. The variations of the family are therefore identical with the functions $\eta_i(x)$ originally prescribed, since when the $\zeta_r(x)$ are given, there is only one set of solutions of equations (11) with given initial values $\eta_i(x_1)$ at $x = x_1$.

Some slight modifications in the existence theorems referred to are required in order to prove the continuity properties of the family $y_i = y_i(x, b)$ described in the theorem. These are due to the fact that the functions $z_i(x)$ defined by the arc E_{12} are continuous but not necessarily differentiable. The results described can be derived without difficulty, however, when the arc E_{12} has no corners. If the arc E_{12} has corners the existence theorems must be applied successively to the x-intervals between the corner-values of x with initial conditions at the beginning of each interval so chosen that the functions $y_i(x, b)$ are continuous.

COROLLARY. *If a matrix*

$$\left\| \begin{matrix} \eta_{11} & \cdots & \eta_{1\mu} \\ \cdots & \cdots & \cdots \\ \eta_{n1} & \cdots & \eta_{n\mu} \end{matrix} \right\|$$

whose columns are μ sets of admissible variations along an admissible arc E_{12}, is given, then there exists a μ-parameter family of admissible arcs $y_i = y_i(x, b_1, \cdots, b_\mu)$ containing E_{12} for the values $b_1 = \cdots = b_\mu = 0$ and having the functions η_{is} $(i = 1, \cdots, n)$ as its variations with respect to b_s along E_{12}. The continuity properties of the family are similar to those described in the preceding theorem.

* Bolza [3, pp. 168 ff.]; Bliss [14, 15].

This is proved as above with the equations

$$\phi_\alpha = 0, \quad \phi_r = z_r(x) + b_1\zeta_{r1} + \cdots + b_\mu\zeta_{r\mu}$$
$$(\alpha = 1, \cdots, m;\ r = m+1, \cdots, n)$$

replacing equations (12).

4. *The first variation of I.* If the functions $y_i(x, b)$ defining a one-parameter family of admissible arcs containing E_{12} for $b = 0$ are substituted in I then I becomes the function of b defined by the formula

$$I(b) = \int_{x_1}^{x_2} f[x, y(x, b), y'(x, b)]\,dx.$$

The derivative of this function with respect to b at the value $b = 0$ is the expression

$$I_1(\eta) = \int_{x_1}^{x_2} (f_{y_i}\eta_i + f_{y_i'}\eta_i')\,dx \tag{13}$$

where i is as agreed an umbral symbol and the arguments of the derivatives of f are the functions $y_i(x)$ defining E_{12}.

The expression $I_1(\eta)$ is called the *first variation* of I along the arc E_{12}. For the proofs of the succeeding sections it is desirable to have another form of it. Let λ_0 be a constant and $\lambda_i(x)$ $(i = 1, \cdots, n)$ functions of x on the interval x_1x_2, and let F be defined by the equation

$$F(x, y, y', \lambda) = \lambda_0 f + \lambda_1\phi_1 + \cdots + \lambda_n\phi_n.$$

Since the variations η, ζ satisfy the equations (11) the value of $\lambda_0 I_1(\eta)$ is not altered if we add the sum $\lambda_\alpha\Phi_\alpha + \lambda_r(\Phi_r - \zeta_r)$ to its integrand. Then we have

$$\lambda_0 I_1(\eta) = \int_{x_1}^{x_2} (F_{y_i}\eta_i + F_{y_i'}\eta_i' - \lambda_r\zeta_r)\,dx. \tag{14}$$

So far the functions $\lambda_i(x)$ have been entirely arbitrary. We now determine them so that the equations

$$F_{y_i'} = \int_{x_1}^{x} F_{y_i}\,dx + c_i \qquad (i = 1, \cdots, n) \tag{15}$$

are satisfied for an arbitrarily selected set of constants λ_0, c_i. This is possible since if we introduce the new variables

$$v_i = F_{y_i'} = \lambda_0 f_{y_i'} + \lambda_1\phi_{1y_i'} + \cdots + \lambda_n\phi_{ny_i'} \qquad (i = 1, \cdots, n) \tag{16}$$

the equations (15) are equivalent to the equations and initial conditions

$$dv_i/dx = F_{y_i} = A_{i1}v_1 + \cdots + A_{in}v_n + B_i, \quad v_i(x_1) = c_i \qquad (i = 1, \cdots, n) \tag{17}$$

the coefficients A, B being found by solving the equations (16) for $\lambda_1, \cdots, \lambda_n$ and substituting in F_{y_i}. The equations (17) have unique solutions $v_i(x)$ which are continuous on the interval x_1x_2 and which have continuous derivatives except possibly at the values of x defining the corners of E_{12} where the coefficients A, B may be discontinuous. Equations (16) then determine uniquely the functions $\lambda_i(x)$ continuous except possibly at the corner values of x.

With the help of equations (15) the expression (14) for $\lambda_0 I_1(\eta)$ now takes the form

$$(18) \qquad \lambda_0 I_1(\eta) = -\int_{x_1}^{x_2} \lambda_r \zeta_r dx - c_i \eta_i(x_1) + \eta_i(x_2) F_{y_i'}(x_2)$$

where $F_{y_i'}(x_2)$ represents the value of $F_{y_i'}$ at $x = x_2$. This auxiliary formula will be useful in the next section.

5. *The Euler-Lagrange multiplier rule.* We are now in a position to deduce the famous multiplier rule giving the differential equations which must be satisfied by a minimizing arc E_{12} for the Lagrange problem. The rule was discussed for a special case by Euler in 1744, and generalized by Lagrange whose proof was exceedingly faulty. One difficulty with Lagrange's proof was overcome by Mayer in 1886, and the proof was finally completed when Kneser in 1900 and Hilbert in 1905 removed the last serious defects.* The proof given here is quite different in some respects from those in the literature and is an extension of them.

Suppose that a matrix whose columns are $2n+1$ sets of admissible variations

$$(19) \qquad \begin{Vmatrix} \eta_{11} & \cdot & \cdot & \cdot & \eta_{1,2n+1} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \eta_{n1} & \cdot & \cdot & \cdot & \eta_{n,2n+1} \end{Vmatrix}$$

is given. We have seen above that there is a $(2n+1)$-parameter family $y_i(x, b_1, \cdots, b_{2n+1})$ of admissible arcs containing E_{12} for $b_1 = \cdots = b_{2n+1} = 0$ and having the columns of the matrix above as its variations. When the functions defining this family are inserted in the integral I that integral becomes a function $I(b_1, \cdots, b_{2n+1})$ which for $b_1 = \cdots = b_{2n+1} = 0$ takes the value I_0 of the integral along the arc E_{12}. If we let $(x_1, y_{11}, \cdots, y_{n1})$ and $(x_2, y_{12}, \cdots, y_{n2})$ represent the two end-points on the arc E_{12} then the equations

* For the details of the objections to Lagrange's proof and an excellent historical sketch see Bolza [3, p. 566].

$$
(20)\qquad
\begin{gathered}
I(b_1, \cdots, b_{2n+1}) = I_0 + u, \\
y_i(x_1, b_1, \cdots, b_{2n+1}) = y_{i1}, \\
y_i(x_2, b_1, \cdots, b_{2n+1}) = y_{i2}, \\
(i = 1, \cdots, n)
\end{gathered}
$$

in the variables $u, b_1, \cdots, b_{2n+1}$ have the initial solution $(u, b_1, \cdots, b_{2n+1}) = (0, 0, \cdots, 0)$. If the functional determinant of the first members of these equations with respect to $b_1, \cdots, b_{2n+1}$ is different from zero at this solution, then well-known implicit function theorems tell us that the equations (20) have solutions not only for $u = 0$ but also for every value of u near $u = 0$. There are therefore arcs in the family $y_i(x, b_1, \cdots, b_{2n+1})$ joining the end-points 1 and 2 of E_{12} and giving I values $I_0 + u$ greater than I_0 when u is positive, and similar arcs giving it values less than I_0 when u is negative, which is impossible if E_{12} is a minimizing arc. Hence the functional determinant of the equations (20) must be zero at $(u, b_1, \cdots, b_{2n+1}) = (0, 0, \cdots, 0)$.

The value of this functional determinant is

$$
(21)\qquad
\left\|
\begin{array}{ccc}
I_1(\eta_1) & \cdots & I_1(\eta_{2n+1}) \\
\eta_{11}(x_1) & \cdots & \eta_{1,2n+1}(x_1) \\
\cdots & \cdots & \cdots \\
\eta_{n1}(x_1) & \cdots & \eta_{n,2n+1}(x_1) \\
\eta_{11}(x_2) & \cdots & \eta_{1,2n+1}(x_2) \\
\cdots & \cdots & \cdots \\
\eta_{n1}(x_2) & \cdots & \eta_{n,2n+1}(x_2)
\end{array}
\right\|
$$

where in the first row only the second subscripts of the η's are indicated. It must vanish for every choice of the matrix (19) of admissible variations. Suppose $p < 2n + 1$ the highest rank attainable for (21) and suppose the matrix (19) chosen so that this rank is actually attained. Let λ_0, c_i, d_i $(i = 1, \cdots, n)$ be a set of constants not all zero satisfying the linear equations whose coefficients are the columns of the determinant (21). Normally the constant λ_0 will be different from zero, but in Section 7 the case $\lambda_0 = 0$ is discussed in more detail. In both cases the equation

$$
\lambda_0 I_1(\eta) + c_i \eta_i(x_1) + d_i \eta_i(x_2) = 0
$$

must be satisfied for every set of admissible variations $\eta_i(x)$ whatsoever, since otherwise by deleting a suitable one of the columns of the determinant (21) and replacing it by a set $I_1(\eta)$, $\eta_i(x_1)$, $\eta_i(x_2)$ which does not satisfy the last equation, the determinant could be made to have the rank $p + 1$. If the first term of the last equation is replaced by its value (18) the equation takes the form

$$-\int_{x_1}^{x_2} \lambda_r \zeta_r dx + \eta_i(x_2)[d_i + F_{y_i'}(x_2)] = 0$$

and it must be satisfied for every choice of the admissible variations $\eta_i(x)$, i. e. for every choice of the functions $\zeta_r(x)$ and the end values $\eta_i(x_2)$, since for every such choice there is a set of admissible variations defined by the equations (11). It follows readily that the conditions

$$\lambda_r(x) \equiv 0, \quad d_i = -F_{y_i'}(x_2) \qquad (r = m+1, \cdots, n;\ i = 1, \cdots, n) \tag{22}$$

must be satisfied. For the set of multipliers λ_0, $\lambda_i(x)$ $(i = 1, \cdots, n)$ for which the equations (15) are satisfied it is evident then that all are identically zero except the first $m+1$. The first $m+1$ of them are not all identically zero, however, since otherwise F would vanish identically and equations (15) and (22) would require the constants c_i, d_i all to be zero as well as λ_0, which we know not to be the case. Hence we have the following theorem:

For every minimizing arc E_{12} there exists a set of constants c_i $(i = 1, \cdots, n)$ and a function

$$F(x, y, y', \lambda) = \lambda_0 f + \lambda_1(x)\phi_1 + \cdots + \lambda_m(x)\phi_m \tag{23}$$

such that the equations

$$F_{y_i'} = \int_{x_1}^{x} F_{y_i} dx + c_i \tag{24}$$

are satisfied at every point of E_{12}. The constant λ_0 and the functions $\lambda_\alpha(x)$ $(\alpha = 1, \cdots, m)$ are not all identically zero on x_1x_2 and are continuous except possibly at values of x defining corners of E_{12}.

This is a modification of the Euler-Lagrange multiplier rule. We get the rule in its classical form by differentiating the equations (24). The two following corollaries are immediate:

COROLLARY I. THE EULER-LAGRANGE MULTIPLIER RULE. *On every sub-arc between corners of a minimizing arc E_{12} the differential equations*

$$\phi_\alpha(x, y, y') = 0, \quad (d/dx)F_{y_i'} = F_{y_i} \quad (\alpha = 1, \cdots, m;\ i = 1, \cdots, n) \tag{25}$$

must be satisfied, where F is the function (23).

COROLLARY II. THE CORNER CONDITION. *At every corner of a minimizing arc E_{12} the conditions*

$$F_{y_i'}[x, y, y'(x-0), \lambda(x-0)] = F_{y_i'}[x, y, y'(x+0), \lambda(x+0)] \qquad (i = 1, \cdots, n) \tag{26}$$

must be satisfied.

Condition (26) is a consequence of the fact that the second member of (24) is continuous at a corner as well as elsewhere.

There is a third consequence of the equations (24) which is also important. If the functions and multipliers belonging to E_{12} are $y_i(x)$, λ_0, $\lambda_\alpha(x)$ then the $n+m$ equations

$$F_{y_i'}[x, y(x), z, \mu] = \int_{x_1}^{x} F_{y_i}[x, y(x), y'(x), \lambda(x)]dx + c_i,$$

$$\phi_\alpha[x, y(x), z] = 0 \qquad (i = 1, \cdots, n;\ \alpha = 1, \cdots, m)$$

have as solutions the $n+m$ functions $z_i = y_i'(x)$, $\mu_\alpha = \lambda_\alpha(x)$. If the functional determinant

$$R = \begin{vmatrix} F_{y_i' y_k'} & \phi_{\alpha y_i'} \\ \phi_{\alpha y_k'} & 0 \end{vmatrix}$$

of the first members of these equations with respect to the variables z_i, μ_α is different from zero at a point of E_{12} then the existence theorems for implicit functions tell us that the solutions $z_i = y_i'(x)$, $\mu_\alpha = \lambda_\alpha(x)$ of the equations have continuous derivatives of as many orders as the equations themselves have continuous partial derivatives in the variables x, z_i, μ_α. Between corners this is at least one, and we have the following third corollary:

COROLLARY III. THE DIFFERENTIABILITY CONDITION. *Near a point of a minimizing arc E_{12} at which the determinant R is different from zero the functions $y_i(x)$ defining E_{12} have continuous second derivatives and the multipliers $\lambda_\alpha(x)$ have continuous first derivatives.*

The proof given above for the Euler-Lagrange multiplier rule is an extension of the ones ordinarily given because the hypothesis (c) Section 1 is less restrictive than usual. The unsymmetrical assumption commonly made is that a particular one of the determinants of the matrix $\| \phi_{\alpha y_i'} \|$ stays different from zero at every point of E_{12}. The enlargement of the system $\phi_\alpha = 0$ to the system (10) is the device which permits the generalization here made. Equations (24) are recent developments which were unknown to Euler and Lagrange and which are not always deduced even in modern presentations of the subject. They justify the useful Corollaries II and III besides the multiplier rule.

6. *The extremals.* An admissible arc and set of multipliers

$$y_i = y_i(x), \quad \lambda_0, \quad \lambda_\alpha = \lambda_\alpha(x) \tag{27}$$
$$(i = 1, \cdots, n;\ \alpha = 1, \cdots, m;\ x_1 \leqq x \leqq x_2)$$

is called an *extremal* if it has continuous derivatives $y_i'(x)$, $y_i''(x)$, $\lambda_\alpha'(x)$

on the interval x_1x_2, and if furthermore it satisfies the Euler-Lagrange equations (25). The minimizing curves for applications of the theory of the calculus of variations are found among the extremals and it is highly desirable, therefore, that we should examine more thoroughly the differential equations defining these curves and determine how large a family the extremals really form. A minimizing curve must always be a solution of the equations (25), even if it has corners or is without the derivatives $y_i''(x)$, $\lambda_a'(x)$ mentioned above, but such minimizing curves are relatively rare.

The most direct way to characterize the family of extremals satisfying equations (25) is to replace these equations by the equivalent system

$$
\begin{aligned}
&(d/dx)F_{y_i'} - F_{y_i} = F_{y_i'x} + F_{y_i'y_k}y_k' + F_{y_i'y_k'}y_k'' + F_{y_i'\lambda_\beta}\lambda_\beta' - F_{y_i} = 0,\\
(28)\qquad &(d/dx)\phi_a = \phi_{ax} + \phi_{ay_k}y_k' + \phi_{ay_k'}y_k'' = 0,\\
&\phi_a\,[x_1, y(x_1), y'(x_1)] = 0.
\end{aligned}
$$

The first two of these equations are linear in the variables y_k'', λ_β' and the determinant of the coefficients of these variables is precisely the determinant R of page 684. Near an extremal E_{12} on which R is different from zero these two equations can therefore be solved for y_k'', λ_β' and they are readily seen to be equivalent to a system

$$
(29)\qquad dy_k/dx = y_k', \quad dy_k'/dx = G_k(x, y, y', \lambda), \quad d\lambda_\beta/dx = H_\beta(x, y, y', \lambda)
$$

in the so-called normal form.* Known existence theorems for differential equations now tell us that an extremal E_{12} along which R is different from zero is a member of a family of solutions of equations (29) depending upon $2n + m$ arbitrary constants, since the number of dependent variables y_k, y_k', λ_β in these equations is $2n + m$. If we impose further the m relations in the third row of equations (28) then m of these constants will be determined as function of the $2n$ others, so that the final result is that an extremal along which R is different from zero is a member of a $2n$-parameter family of extremals satisfying equations (25).

For theoretical purposes the properties of the $2n$-parameter family of extremals may be determined most conveniently by a second method.† For the purpose of introducing n new variables v_i and eliminating the $n + m$ variables y_i', λ_a let us consider the system of $n + m$ equations

$$
(30)\qquad F_{y_i'}(x, y, y', \lambda) = v_i, \quad \phi_a(x, y, y') = 0.
$$

The functional determinant of the first members of these equations with respect

* Bolza [3, p. 589].

† Bolza [3, p. 590].

to the variables y_k', λ_β is again the determinant R of page 684. Known theorems on implicit functions tell us then that near an extremal E_{12} on which R is different from zero the equations (30) have solutions

$$(31) \qquad y_k' = \Psi_k(x, y, v), \quad \lambda_\beta = \Pi_\beta(x, y, v)$$

possessing continuous partial derivatives of the first three orders since the first members of equations (30) have such derivatives. The system of equations (25) is now equivalent to the system in normal form

$$(32) \quad dy_k/dx = \Psi_k(x, y, v), \quad dv_k/dx = F_{y_k}[x, y, \Psi(x, y, v), \quad \Pi(x, y, v)]$$

in the variables x, y_k, v_k. Evidently every solution $y_k(x)$, $\lambda_\beta(x)$ of equations (25) defines a set of functions $v_k(x)$ satisfying equations (30) and (31), and therefore also the system (32). Conversely every solution $y_k(x)$, $v_k(x)$ of equations (32) defines a set of functions $\lambda_\beta(x)$ by means of equations (31) with which it satisfies equations (30), and therefore also the original system (25).

Through every initial element

$$(x_0, y_0, v_0) = (x_0, y_{10}, \cdots, y_{n0}, v_{10}, \cdots, v_{n0})$$

in a neighborhood of the set of values (x, y, v) on the extremal E_{12} there passes a unique solution

$$(33) \qquad y_i = y_i(x, x_0, y_0, v_0), \quad v_i = v_i(x, x_0, y_0, v_0)$$

of the equations (32) for which the functions y_i, y_{ix}, v_i, v_{ix} have continuous partial derivatives of the first three orders since the second members of equations (32) have such derivatives. The equations expressing the fact that the solutions (33) passes through (x_0, y_0, v_0) are

$$y_{i0} = y_i(x_0, x_0, y_0, v_0), \quad v_{i0} = v_i(x_0, x_0, y_0, v_0),$$

and from them we find

$$(34) \quad \begin{array}{ll} \delta_{ik} = (\partial/\partial y_{k0}) y_i(x_0, x_0, y_0, v_0), & 0 = (\partial/\partial v_{k0}) y_i(x_0, x_0, y_0, v_0), \\ 0 = (\partial/\partial y_{k0}) v_i(x_0, x_0, y_0, v_0), & \delta_{ik} = (\partial/\partial v_{k0}) v_i(x_0, x_0, y_0, v_0), \end{array}$$

where δ_{ik} is 1 or 0 when $k = i$ or $k \neq i$, respectively. Since every curve of this system (33) has on it an initial element for which $x = x_1$ we lose none of the curves if we replace x_0 by the fixed value x_1. Let us for convenience rename the constants y_{i0}, v_{i0} and call them a_i, b_i respectively. Then the family (33) takes the form

$$(35) \qquad y_i = y_i(x, a, b), \quad v_i = v_i(x, a, b)$$

and it follows readily from equations (34) that the determinant

$$
(36) \qquad \begin{vmatrix} \dfrac{\partial y_i}{\partial a_k} & \dfrac{\partial y_i}{\partial b_k} \\ \dfrac{\partial v_i}{\partial a_k} & \dfrac{\partial v_i}{\partial b_k} \end{vmatrix}
$$

has the value 1 at $x=x_1$. When we substitute the functions (35) in equations (31) a set of functions $\lambda_a(x,a,b)$ is determined, and we have the final result:

Every extremal E_{12} along which the determinant R is different from zero is a member of a 2n-parameter family of extremals

$$
(37) \qquad y_i = y_i(x,a,b), \quad \lambda_a = \lambda_a(x,a,b)
$$

for special values a_0, b_0 of the parameters. The functions y_i, y_{ix}, v_i, v_{ix}, λ_β have continuous partial derivatives of the first three orders in a neighborhood of the values (x,a,b) defining E_{12}, and at the special values (x_1,a_0,b_0) the determinant (36) *is different from zero.*

Thus again we have established the existence of a family of extremals containing $2n$ arbitrary constants.

7. *Normal admissible arcs.* An admissible arc $y_i=y_i(x)$ $(x_1 \leqq x \leqq x_2)$ is said to be *normal* if there exist for it $2n$ sets of admissible variations for which the determinant

$$
(38) \qquad \begin{vmatrix} \eta_{11}(x_1) & \cdot & \cdot & \cdot & \eta_{1,2n}(x_1) \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \eta_{n1}(x_1) & \cdot & \cdot & \cdot & \eta_{n,2n}(x_1) \\ \eta_{11}(x_2) & \cdot & \cdot & \cdot & \eta_{1,2n}(x_2) \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \eta_{n1}(x_2) & \cdot & \cdot & \cdot & \eta_{n,2n}(x_2) \end{vmatrix}
$$

is different from zero. It is *normal on a sub-interval* $\xi_1\xi_2$ of x_1x_2 if there exist $2n$ sets of admissible variations for which the last determinant is different from zero when x_1 is replaced by ξ_1 and x_2 by ξ_2. In the sequel we shall frequently need to restrict our proofs to arcs which are *normal on every sub-interval of* x_1x_2.

These definitions doubtless seem at first sight somewhat artificial. If an admissible arc E_{12} is not normal, however, it is in general true that no other admissible arcs near it pass through the end points 1 and 2 of E_{12}, and hence that near E_{12} the class of arcs in which we seek to minimize the integral I has in it only E_{12} itself. The minimum problem in such a case would not be

of interest. We shall presently see that there are always an infinity of admissible arcs through the ends of E_{12} when E_{12} is normal.

A necessary and sufficient condition that an admissible arc be normal is that there exists for it no set of multipliers λ_0, $\lambda_\alpha(x)$ having $\lambda_0 = 0$ with which it satisfies the equations:

$$F_{y_i'} = \int_{x_1}^{x} F_{y_i} dx + c_i.$$

For a normal extremal arc multipliers in the form $\lambda_0 = 1$, $\lambda_\alpha(x)$ always exist and in this form they are unique.

The processes of Section 5 show that an admissible arc which is not normal has surely a set of multipliers with $\lambda_0 = 0$, since the linear equations whose coefficients are the columns of the determinant (21) have for such an arc a set of solutions λ_0, c_i, d_i with $\lambda_0 = 0$. The first sentence of the theorem will then be justified if we can show that a normal admissible arc has no set of multipliers with $\lambda_0 = 0$.

Suppose that there were a normal admissible arc with a set of multipliers having $\lambda_0 = 0$. Its function F would have the form

$$F = \lambda_1 \phi_1 + \cdots \lambda_m \phi_m$$

and every set of admissible variations along it would satisfy the equation

$$0 = \int_{x_1}^{x_2} \lambda_\alpha \Phi_\alpha dx = \int_{x_1}^{x_2} (F_{y_i} \eta_i + F_{y_i'} \eta_i') dx = F_{y_i'}(x_2)\eta_i(x_2) - F_{y_i'}(x_1)\eta_i(x_1)$$

on account of the equations of variations (9) and the equations of the theorem above. Since there is a determinant (38) different from zero it follows that the derivatives $F_{y_i'}$ would all vanish at x_1 and x_2 on our extremal. If we define the variables v_i again by equations (30), or by equations (16) with $\lambda_0 = \lambda_{m+1} = \cdots = \lambda_n = 0$, then in equations (17) the coefficients B_i and the initial values $v_i(x_1) = F_{y_i'}(x_1)$ would all vanish. The only continuous solutions of equations (17) under these circumstances are the functions $v_i(x) \equiv 0$, and equations (16) then imply that the multipliers $\lambda_\alpha(x)$ would all vanish identically, which is not the case. Hence a normal admissible arc can not have a set of multipliers with constant multiplier λ_0 equal to zero.

When an extremal arc has multipliers with $\lambda_0 \neq 0$ the multipliers can evidently all be divided by λ_0 to obtain a set of the form $\lambda_0 = 1$, $\lambda_\alpha(x)$. If there were a second set $\lambda_0 = 1$, $\Lambda_\alpha(x)$ the differences 0, $\Lambda_\alpha - \lambda_\alpha$ would also be a set of multipliers for E_{12} with the constant multiplier zero. We have just seen that this is impossible for a normal extremal unless $\Lambda_\alpha - \lambda_\alpha \equiv 0$, so that the multipliers $\lambda_0 = 1$, $\lambda_\alpha(x)$ of a normal extremal E_{12} are unique.

In every neighborhood of a normal admissible arc E_{12} there are an infinity of admissible arcs with the same end-points 1 *and* 2.

To prove this consider the set of $2n$ admissible variations for E_{12} appearing in the determinant (38) and an additional set $\eta_i(x)$. From the results of Section 3 we know that there is a family of admissible arcs $y_i = Y_i(x, b, b_1, b_2, \cdots, b_{2n})$ containing E_{12} when $b = b_1 = \cdots = b_{2n} = 0$ and having the sets $\eta_i(x)$, $\eta_{is}(x)$ $(s = 1, \cdots, 2n)$ as its variations. The $2n$ equations

$$(39) \qquad Y_i(x_1, b, b_1, \cdots, b_{2n}) = y_{i1}, \quad Y_i(x_2, b, b_1, \cdots, b_{2n}) = y_{i2}$$

have the initial solution $(b, b_1, \cdots, b_{2n}) = (0, 0, \cdots, 0)$ at which the functional determinant of their first members with respect to $b_1, \cdots, b_{2n}$ is the determinant (38) and different from zero. Hence by the usual implicit function theorems these equations have solutions $b_s = B_s(b)$ $(s = 1, \cdots, 2n)$ with initial values $B_s(0) = 0$, and the one parameter family of admissible arcs

$$(40) \qquad y_i = Y_i[x, b, B_1(b), \cdots, B_{2n}(b)] = y_i(x, b)$$

defined by them contains the extremal E_{12} for $b = 0$ and has all its curves passing through the points 1 and 2.

COROLLARY. *If each function $\eta_i(x)$ of a set of admissible variations for a normal admissible arc E_{12} vanishes at x_1 and x_2 then there is a one-parameter family of admissible arcs $y_i = y_i(x, b)$ passing through the points* 1 *and* 2, *containing E_{12} for the parameter value $b = 0$, and having the set $\eta_i(x)$ as its variations along E_{12}.*

Let us suppose that in the construction of the family (40) the set $\eta_i(x)$ of the Corollary has been used. Since these functions all vanish at x_1 and x_2 we find from equations (39), by differentiating with respect to b and setting $b = 0$, that

$$\eta_{is}(x_1) B_s'(0) = 0, \quad \eta_{is}(x_2) B_s'(0) = 0.$$

Since the determinant (38) is different from zero these imply that all the derivatives $B_s'(0)$ vanish. Hence the family (40) has the variations

$$y_{ib}(x, 0) = \eta_i(x) + Y_{ib_s} B_s'(0) = \eta_i(x).$$

We know already that the family contains E_{12} for $b = 0$ and has all of its curves passing through 1 and 2.

8. *Problems with variable end-points.** It happens that a number of important applications of the theory of the Lagrange problem are of a slightly

* See Bliss [16].

2

different type from that described in Section 1. In order to include them as special cases we must permit variable end-points for the curves of the class in which we are seeking a minimum for I. We shall endeavor to find among the arcs

$$y_i = y_i(x) \qquad (i = 1, \cdots, n;\ x_1 \leqq x \leqq x_2)$$

satisfying the system of equations

$$\phi_\alpha(x, y, y') = 0 \qquad (\alpha = 1, \cdots, m < n)$$

and having end-points satisfying the equations

$$\psi_\mu[x_1, y(x_1), x_2, y(x_2)] = 0 \qquad (\mu = 1, \cdots, p \leqq 2n + 2) \tag{41}$$

one which minimizes the integral I. The number p must not exceed the number $2n + 2$ of end values x_1, y_{i1}, x_2, y_{i2} since otherwise equations (41) would in general have no solutions. The problem of Section 1 is a special case of this one with the system (41) having the special form

$$x_1 - \alpha_1 = y_{i1} - \beta_{i1} = x_2 - \alpha_2 = y_{i2} - \beta_{i2} = 0$$

for which p has exactly the value $2n + 2$.

Suppose now that E_{12} is a minimizing arc for the new problem with end values $(x_1, y_{i1}, x_2, y_{i2})$. We add to the hypotheses (a), (b), (c) of Section 1 the assumption

(d) the functions ψ_μ have continuous derivatives up to and including those of the fourth order near the end-values $(x_1, y_{i1}, x_2, y_{i2})$ of E_{12}, and at these values the $p \times (2n + 2)$-dimensional matrix

$$\| \psi_{\mu x_1} \quad \psi_{\mu y_{i1}} \quad \psi_{\mu x_2} \quad \psi_{\mu y_{i2}} \| \tag{42}$$

has rank p.

The last part of this assumption implies that the equations $\psi_\mu = 0$ are all independent.

It is evident that the arc E_{12} must minimize I in the class of admissible arcs having the same end-values, and we can infer at once that it must have a system of multipliers with which it satisfies the necessary conditions deduced in Section 5. But it is important that we should analyse the situation somewhat more closely. Let

$$y_i = y_i(x, b) \quad [x_1(b) \leqq x \leqq x_2(b)] \tag{43}$$

be a one-parameter family of admissible arcs containing E_{12} for $b = 0$ whose end-values satisfy the equations

$$\psi_\mu\{x_1(b), y_i[x_1(b), b], x_2(b), y_i[x_2(b), b]\} = 0.$$

If we use the notations $x_{1b}(0) = \xi_1$, $x_{2b}(0) = \xi_2$ the derivatives of these equations with respect to b for $b = 0$ are the system

$$\text{(44)} \qquad \begin{aligned}\Psi_\mu(\xi, \eta) = (\psi_{\mu x_1} + \psi_{\mu y_{i1}} y'_{i1})\xi_1 + \psi_{\mu y_{i1}}\eta_i(x_1) \\ + (\psi_{\mu x_2} + \psi_{\mu y_{i2}} y'_{i2})\xi_2 + \psi_{\mu y_{i2}}\eta_i(x_2).\end{aligned}$$

These are the *equations of variation* on E_{12} for the functions ψ_μ. When the family (43) is substituted in the integral I we find for the first variation the formula

$$I_1(\xi, \eta) = \int_{x_1}^{x_2} (f_{y_i}\eta_i + f_{y_i'}\eta_i')\,dx + f(x_2)\xi_2 - f(x_1)\xi_1$$

where $f(x_1)$ and $f(x_2)$ are the values of f at the points 1 and 2 on E_{12}. With the help of the expression (18) we may also write

$$\text{(45)} \qquad \begin{aligned}\lambda_0 I_1(\xi, \eta) = -\int_{x_1}^{x_2} \lambda_r \zeta_r\,dx - \lambda_0 f(x_1)\xi_1 \\ - c_i\eta_i(x_1) + \lambda_0 f(x_2)\xi_2 + \eta_i(x_2)F_{y_i'}(x_2)\end{aligned}$$

where the constants c_i may be arbitrarily chosen.

A set of admissible variations for the present problem is a set ξ_1, ξ_2, $\eta_i(x)$ in which ξ_1 and ξ_2 are arbitrary constants and the functions $\eta_i(x)$ form a set of admissible variations in the sense of Section 3. For a matrix

$$\left\| \begin{matrix} \xi_{11} & \cdot & \cdot & \cdot & \xi_{1,p+1} \\ \xi_{21} & \cdot & \cdot & \cdot & \xi_{2,p+1} \\ \eta_{11} & \cdot & \cdot & \cdot & \eta_{1,p+1} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \eta_{n1} & \cdot & \cdot & \cdot & \eta_{n,p+1} \end{matrix} \right\|$$

whose columns are sets of admissible variations there exists a family

$$\text{(46)} \qquad \begin{aligned} y_i = y_i(x, b_1, \cdots, b_{p+1}) \\ x_1(b_1, \cdots, b_{p+1}) \leqq x \leqq x_2(b_1, \cdots, b_{p+1}) \end{aligned}$$

containing E_{12} for $(b_1, \cdots, b_{p+1}) = (0, \cdots, 0)$ and having the sets $\xi_{1\sigma}$, $\xi_{2\sigma}$, $\eta_{i\sigma}(x)$ $(\sigma = 1, \cdots, p+1)$ as its variations along E_{12} with respect to the parameters b_σ. Such a family is that of the Corollary on page 679 with the functions

$$x_\rho(b_1, \cdots, b_{p+1}) = x_\rho + b_\sigma \xi_{\rho\sigma}, \qquad (\rho = 1, 2)$$

adjoined. When the equations of the family (46) are substituted in the integral I and the functions ψ_μ, these become functions of $b_1, \cdots, b_{p+1}$. The first members of the equations

$$I(b_1, \cdots, b_{p+1}) = I_0 + u,$$
$$\psi_\mu(b_1, \cdots, b_{p+1}) = 0$$

must have their functional determinant equal to zero for $(b_1, \cdots, b_{p+1}) = (0, \cdots, 0)$ by the same argument as that on page 682. This determinant is

$$(47) \qquad \begin{vmatrix} I_1(\xi_1, \eta_1) & \cdots & I_1(\xi_{p+1}, \eta_{p+1}) \\ \Psi_1(\xi_1, \eta_1) & \cdots & \Psi_1(\xi_{p+1}, \eta_{p+1}) \\ \cdots & \cdots & \cdots \\ \Psi_p(\xi_1, \eta_1) & \cdots & \Psi_p(\xi_{p+1}, \eta_{p+1}) \end{vmatrix}$$

in which only the second subscripts of the sets $\xi_{1\sigma}$, $\xi_{2\sigma}$, $\eta_{i\sigma}$ have been indicated. From its vanishing we argue as on page 682 that there exists a set of constants λ_0, $d_1, \cdots, d_p$ not all zero such that the equation

$$\lambda_0 I_1(\xi, \eta) + d_\mu \Psi_\mu(\xi, \eta) = 0$$

must hold for every set of admissible variations ξ_1, ξ_2, $\eta_i(x)$. With the help of formulas (44) and (45) this becomes

$$\begin{aligned} -\int_{x_1}^{x_2} \lambda_r \zeta_r dx &+ [-\lambda_0 f(x_1) + d_\mu(\psi_{\mu x_1} + \psi_{\mu y_{i1}} y'_{i1})]\, \xi_1 \\ &+ [\lambda_0 f(x_2) + d_\mu(\psi_{\mu x_2} + \psi_{\mu y_{i2}} y'_{i2})]\, \xi_2 \\ &+ [-c_i + d_\mu \psi_{\mu y_{i1}}]\, \eta_i(x_1) \\ &+ [F_{y_i'}(x_2) + d_\mu \psi_{\mu y_{i2}}]\, \eta_i(x_2) = 0. \end{aligned}$$

After the arbitrary constants c_i in (45) have been so chosen that the coefficients of the terms in $\eta_i(x_1)$ in the last expression all vanish it follows by an argument like that of page 683 that $\lambda_{\mu+1} \equiv \cdots \equiv \lambda_n \equiv 0$ and that the coefficients of ξ_1, ξ_2, $\eta_i(x_2)$ also vanish. This result is equivalent to saying that all the determinants of order $p+1$ of the matrix

$$\begin{Vmatrix} -\lambda_0 f(x_1) & -F_{y_i'}(x_1) & \lambda_0 f(x_2) & F_{y_i'}(x_2) \\ \psi_{\mu x_1} + \psi_{\mu y_{i1}} y_{i1}' & \psi_{\mu y_{i1}} & \psi_{\mu x_2} + \psi_{\mu y_{i2}} y_{i2}' & \psi_{\mu y_{i2}} \end{Vmatrix}$$

are zero, since the constants c_i are from equations (15) the values $F_{y_i'}(x_1)$, and since the multipliers $1, d_1, \cdots, d_p$ satisfy all the linear equations whose coefficients are columns of the matrix. The rank of the last matrix is unchanged when one column is multiplied by a factor and added to another, and $\lambda_0 f = F$ on the admissible arc E_{12}, so that these results can be formulated as follows:

For every minimizing arc for the problem of Lagrange with variable end-points there exists a set of constants c_i $(i = 1, \cdots, n)$ *and a function*

$$F(x, y, y', \lambda) = \lambda_0 f + \lambda_1(x)\phi_1 + \cdots + \lambda_m(x)\phi_m$$

such that the equations

$$F_{y_i'} = \int_{x_1}^{x} F_{y_i} dx + c_i$$

are satisfied at every point of E_{12}. *The constant* λ_0 *and the functions* $\lambda_\alpha(x)$ $(\alpha = 1, \cdots, m)$ *are not all identically zero on* x_1x_2 *and are continuous except possibly at values of* x *defining corners of* E_{12}. *Furthermore the end-values of* E_{12} *must be such that all the determinants of order* $p+1$ *of the matrix*

$$(48) \qquad \left\| \begin{matrix} -F(x_1) + y_{i1}'F_{y_i'}(x_1) & -F_{y_i'}(x_1) & F(x_2) - y_{i2}'F_{y_i'}(x_2) & F_{y_i'}(x_2) \\ \psi_{\mu x_1} & \psi_{\mu y_{i1}} & \psi_{\mu x_2} & \psi_{\mu y_{i2}} \end{matrix} \right\|$$

are zero. These last conditions are the so-called transversality conditions.

It is clear that the multipliers λ_0, $\lambda_\alpha(x)$ can not all vanish identically on x_1x_2. Otherwise the constants $d_1, \cdots, d_p$ would have to satisfy the linear equations whose coefficients are the columns of the matrix (42) which has rank p. The constants λ_0, $d_1, \cdots, d_p$ would then all be zero which is not the case.

9. *Normal admissible arcs for problems with variable end-points.* A *normal admissible arc* for the problem of Lagrange with variable end-points is one for which there exist p sets of admissible variations $\xi_{1\mu}$, $\xi_{2\mu}$, $\eta_{i\mu}(x)$ $(\mu = 1, \cdots, p)$ such that the matrix

$$(49) \qquad \left| \begin{matrix} \Psi_1(\xi_1, \eta_1) & \cdot\ \cdot\ \cdot & \Psi_1(\xi_p, \eta_p) \\ \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot \\ \Psi_p(\xi_1, \eta_1) & \cdot\ \cdot\ \cdot & \Psi_p(\xi_p, \eta_p) \end{matrix} \right|$$

is different from zero. In the elements of the matrix only the second subscripts of the sets $\xi_{1\mu}$, $\xi_{2\mu}$, $\eta_{i\mu}(x)$ are indicated.

A necessary and sufficient condition that an admissible arc for the problem of Lagrange with variable end-points be normal is that there exists for it no set of multipliers λ_0, $\lambda_\alpha(x)$ *having* $\lambda_0 = 0$ *with which it satisfies the conditions of the last theorem. For a normal extremal arc satisfying the conditions of the last theorem multipliers in the form* $\lambda_0 = 1$, $\lambda_\alpha(x)$ *always exist and in this form they are unique.*

The proof of Section 8 shows that an admissible arc which is not normal has surely a set of multipliers with $\lambda_0 = 0$, since the linear equations whose coefficients are the columns of the determinant (47) have for such an arc solutions λ_0, $d_1, \cdots, d_p$ with $\lambda_0 = 0$.

Suppose now that there were a normal admissible arc satisfying the conditions of the theorem of Section 8 and having $\lambda_0 = 0$. Since the matrix preceding (48) is of rank less than $p+1$ we should then have constants d_μ $(\mu = 1, \cdots, p)$ such that

$$
\begin{aligned}
F(x_1) &= d_\mu(\psi_{\mu x_1} + \psi_{\mu y_{i1}} y'_{i1}), \\
F_{y_i'}(x_1) &= d_\mu \psi_{\mu y_{i1}}, \\
-F(x_2) &= d_\mu(\psi_{\mu x_2} + \psi_{\mu y_{i2}} y'_{i2}), \\
-F_{y_i'}(x_2) &= d_\mu \psi_{\mu y_{i2}}.
\end{aligned} \tag{50}
$$

The numbers $F(x_1)$, $F(x_2)$ would be zero since $\lambda_0 = 0$ and along an admissible arc $F = \lambda_0 f$. After multiplying these equations respectively by ξ_1, $\eta_i(x_1)$, ξ_2, $\eta_i(x_2)$ and adding we should have

$$\eta_i(x_1) F_{y_i'}(x_1) - \eta_i(x_2) F_{y_i'}(x_2) = d_\mu \Psi_\mu(\xi, \eta).$$

The first member of this equation would vanish for every set of admissible variations $\eta_i(x)$, as was proved in Section 7, page 688, and the second member would necessarily have the same property. Since there is a determinant (49) different from zero we should then have $d_\mu = 0$ for every μ, and equations (50) show that $F_{y_i'}(x_1)$ and $F_{y_i'}(x_2)$ would all vanish. As in Section 7, page 688, this would necessitate the vanishing of λ_0, $\lambda_a(x)$ which is impossible. The proof of the uniqueness of the multipliers $\lambda_0 = 1$, $\lambda_a(x)$ is precisely that of Section 7.

In every neighborhood of a normal admissible arc E_{12} for the Lagrange problem with variable end-points there is an infinity of admissible arcs satisfying the end conditions $\psi_\mu = 0$.

The proof is similar to that of the corresponding theorem in Section 7. Select arbitrarily an admissible set of variations ξ_1, ξ_2, $\eta_i(x)$ and p other such sets $\xi_{1\mu}$, $\xi_{2\mu}$, $\eta_{i\mu}(x)$ with determinant (49) different from zero. There is a $p+1$-parameter family

$$
\begin{gathered}
y_i = Y_i(x, b, b_1, \cdots, b_p) \\
X_1(b, b_1, \cdots b_p) \leqq x \leqq X_2(b, b_1, \cdots, b_p)
\end{gathered} \tag{51}
$$

of admissible arcs containing E_{12} for $(b, b_1, \cdots, b_p) = (0, 0, \cdots, 0)$ and having the sets ξ_1, ξ_2, $\eta_i(x)$ and $\xi_{1\mu}$, $\xi_{2\mu}$, $\eta_{i\mu}(x)$ as its variations along E_{12}. The existence of the functions Y_i is a consequence of the corollary of Section 3 above, and we may take $X_\rho = x_\rho + b\xi_\rho + b_\mu \xi_{\rho\mu}$ $(\rho = 1, 2)$. Each function ψ_μ becomes a function $\psi_\mu(b, b_1, \cdots, b_p)$ when the functions (51) defining these arcs are substituted. The equations

$$\psi_\mu(b, b_1, \cdots, b_p) = 0 \tag{52}$$

have the initial solution $(b, b_1, \cdots b_\mu) = (0, 0, \cdots, 0)$ at which the functional determinant of their first members with respect to $b_1, \cdots, b_\mu$ is the determinant (49) different from zero. Hence these equations have p solutions $b_\mu = B_\mu(b)$ with initial values $B_\mu(0) = 0$. The one-parameter family

$$y_i = Y_i[x, b, B_1(b), \cdots, B_p(b)] = y_i(x, b) \qquad x_1(b) \leqq x \leqq x_2(b) \tag{53}$$

where

$$x_\rho(b) = X_\rho[b, B_1(b), \cdots, B_p(b)] \qquad (\rho = 1, 2)$$

contains E_{12} for $b = 0$ and satisfies the equations $\psi_\mu = 0$.

COROLLARY. *If a set of admissible variations ξ_1, ξ_2, $\eta_i(x)$ for a normal admissible arc E_{12} for the Lagrange problem with variable end-points satisfies the equations $\Psi_\mu(\xi, \eta) = 0$, then there exists a one parameter family*

$$y_i = y_i(x, b), \quad x_1(b) \leq x \leq x_2(b)$$

of admissible arcs satisfying the end-conditions $\psi_\mu = 0$, containing E_{12} for the parameter value $b = 0$, and having the set ξ_1, ξ_2, $\eta_i(x)$ as its variations along E_{12}.

If the set ξ_1, ξ_2, $\eta_i(x)$ of the Corollary is used in the construction of the family (53) then we find, by differentiating equations (52) with respect to b and setting $b = 0$, that

$$\Psi_\mu(\xi, \eta) + \Psi_\mu(\xi_\nu, \eta_\nu) B_\nu'(0) = 0.$$

But since the first terms in these equations vanish, and since the determinant (49) is different from zero, it follows that $B_\mu'(0) = 0$ for every μ. Hence the variations of the family (53) are the functions

$$\begin{aligned} y_{ib}(x, 0) &= \eta_i(x) + Y_{ib_\mu} B_\mu'(0) = \eta_i(x), \\ x_{\rho b}(0) &= \xi_\rho + X_{\rho b_\mu} B_\mu'(0) = \xi_\rho, \qquad (\rho = 1, 2) \end{aligned}$$

as required in the Corollary.

CHAPTER II.

APPLICATIONS OF THE EULER-LAGRANGE MULTIPLIER RULE.

10. *The brachistochrone in a resisting medium.* Analytically the problem of the brachistochrone in a plane and in a resisting medium is, as we have seen in Section 2, that of finding among the arcs

$$y = y(x), \quad v = v(x) \quad (x_1 \leq x \leq x_2)$$

satisfying the conditions

$$
\begin{gathered}
vv' - gy' + R(v)(1 + y'^2)^{1/2} = 0, \\
x_1 - \alpha_1 = y_1 - \beta_1 = v_1 - \gamma = x_2 - \alpha_2 = y_2 - \beta_2 = 0,
\end{gathered} \tag{54}
$$

one which minimizes the integral

$$I = \int_{x_1}^{x_2} (1/v)(1 + y'^2)^{1/2} dx.$$

In these expressions primes denote derivatives with respect to x. To apply the Euler-Lagrange rule and the transversality conditions of Section 8 we construct the function

$$
\begin{aligned}
F &= (1/v)(1 + y'^2)^{1/2} + \lambda[vv' - gy' + R(1 + y'^2)^{1/2}] \\
&= H(1 + y'^2)^{1/2} + \lambda(vv' - gy')
\end{aligned}
$$

where H is a convenient symbol * for the expression

$$H = (1/v) + \lambda R(v). \tag{55}$$

The differential equations of the normal extremals are then easily found to be

$$H(dy/ds) = \lambda g + a, \quad v(d\lambda/ds) = H_v, \quad v(dv/ds) = g(dy/ds) - R \tag{56}$$

where s is the length of arc defined by the equation

$$ds = (1 + y'^2)^{1/2} dx$$

and a is a new constant of integration. By eliminating dy and ds from equations (56) we find

$$H(H_v dv + R d\lambda) = (g\lambda + a) g d\lambda,$$

which gives at once, since $H_\lambda = R$, the relation

$$H^2 = (g\lambda + a)^2 + b^2 \tag{57}$$

where b is a second constant of integration. The constant can be taken squared since the first equation (56) shows that H^2 is always greater than $(\lambda g + a)^2$.

Equations (56) and (57) give further

$$
\begin{gathered}
\frac{dy}{dv} = \frac{dy}{ds}\,\frac{ds}{dv} = \frac{v(\lambda g + a)}{g(\lambda g + a) - RH}, \\
\frac{dx}{dv} = \frac{dx}{ds}\,\frac{ds}{dv} = \left[1 - \left(\frac{dy}{ds}\right)^2\right]^{1/2} \frac{ds}{dv} = \frac{bv}{g(\lambda g + a) - RH}.
\end{gathered} \tag{58}
$$

* Bolza [3, p. 577].

Equation (57) is quadratic in λ and when its solution $\lambda = \lambda(v, a, b)$ is substituted in the last equations the values of x and y may be found by quadratures in the form

$$x = \phi(v, a, b) + c, \quad y = \psi(v, a, b) + d, \tag{59}$$

where c and d are again constants of integration. These are the equations of the minimizing arc in parametric form.

It is very easy to set up the matrix (48) for our function F and the five end conditions. It is a square matrix with six rows and columns and its vanishing prescribes the single condition $\lambda(x_2) v(x_2) = 0$. From the equation (57) multiplied by v^2 and equation (55) we then find at $x = x_2$ that $v_2{}^2(a^2 + b^2) = 1$. For the determination of v_2 and the four constants of integration in equations (59) we have therefore in accordance with conditions (54) the five equations

$$\begin{gathered} \phi(v_1, a, b) + c = \alpha_1, \qquad \phi(v_2, a, b) + c = \alpha_2, \\ \psi(v_1, a, b) + d = \beta_1, \qquad \psi(v_2, a, b) + d = \beta_2, \\ v_2{}^2(a^2 + b^2) = 1. \end{gathered} \tag{60}$$

If the resistance function $R(v)$ were known we should now have in equations (57), (56), and (60) the mathematical mechanism for determining possible normal minimizing curves. The adjective possible is used here because the conditions deduced so far have only been shown to be necessary for a normal minimizing arc. They have not been proved to be sufficient to insure a minimum.

11. *Parametric problems in space.* Let us now consider space curves whose equations are given in the parametric form

$$x = x(s), \quad y = y(s), \quad z = z(s) \qquad (s_1 \leq s \leq s_2). \tag{61}$$

The problem to be studied is that of finding among the arcs of this type which satisfy the equation

$$x'^2 + y'^2 + z'^2 - 1 = 0 \tag{62}$$

and join two given points 1 and 2 in xyz-space, one which minimizes an integral of the form

$$I = \int_{s_1}^{s_2} f(x, y, z, x', y', z')\, ds.$$

Primes now denote differentiation with respect to s. Equation (62) restricts the parameter s to be the length of arc measured along the curve (61). If

we agree to measure this length always from the point 1 then the conditions for the curve (61) to pass through 1 and 2 are

$$s_1 = x_1 - \alpha_1 = y_1 - \beta_1 = z_1 - \gamma_1 = x_2 - \alpha_2 = y_2 - \beta_2 = z_2 - \gamma_2 = 0$$

where $(\alpha_1, \beta_1, \gamma_1)$ and $(\alpha_2, \beta_2, \gamma_2)$ are the coördinates of these points. Evidently our problem is one with a variable end-point in $sxyz$-space since s_2 is undetermined.

The function F for normal minimizing arcs is

$$F = f + (\lambda/2)(x'^2 + y'^2 + z'^2 - 1)$$

and the differential equations determining such arcs are

$$(63)\qquad \begin{aligned} &f_x - (d/ds)f_{x'} - \lambda' x' - \lambda x'' = 0, \\ &f_y - (d/ds)f_{y'} - \lambda' y' - \lambda y'' = 0, \\ &f_z - (d/ds)f_{z'} - \lambda' z' - \lambda z'' = 0, \\ &x'^2 + y'^2 + z'^2 = 1. \end{aligned}$$

The sum of the first three of these multiplied, respectively, by x', y', z' gives, with the help of the last one,

$$(64)\qquad (d/ds)(f - x'f_{x'} - y'f_{y'} - z'f_{z'} - \lambda) = 0.$$

The matrix (48) for this problem has eight rows and columns and the vanishing of its determinant demands that at the value s_2

$$(65)\qquad \lambda = f - x'f_{x'} - y'f_{y'} - z'f_{z'}.$$

On account of equation (64) this must be an identity in s.

A very important case is the one for which the function f is positively homogeneous and of the first order in x', y', z', i. e. the one for which the equation

$$(66)\qquad f(x, y, z. kx', ky', kz') = kf(x, y, z, x', y', z')$$

is an identity in its arguments for all $k > 0$. The integral I then has the same value for all parametric representations of the arc (61). The integrands of the length integral and of many other integrals important in the applications of the theory of the Lagrange problem satisfy this condition. When equation (66) is differentiated for k, and the substitution $k = 1$ afterward made, we find the identity

$$(67)\qquad x'f_{x'} + y'f_{y'} + z'f_{z'} = f.$$

From equation (65) it is evident that in this case $\lambda = 0$ and equations (63) become

(68) $f_x - (d/ds)f_{x'} = 0, \quad f_y - (d/ds)f_{y'} = 0, \quad f_z - (d/ds)f_{z'} = 0,$

(69) $$x'^2 + y'^2 + z'^2 - 1 = 0.$$

Only three of these can be independent, since one finds readily that

$$x'P + y'Q + z'R = (d/ds)(f - x'f_{x'} - y'f_{y'} - z'f_{z'}) = 0$$

where P, Q, R are symbols for the first members of equations (68).

12. *Isoperimetric problems.* Suppose that we seek to find in the class of arcs

$$y = y(x) \qquad (x_1 \leq x \leq x_2)$$

joining two given points and satisfying relations of the form

(70) $$\int_{x_1}^{x_2} g_i(x, y, y')\,dx = l_i \qquad (i = 1, \cdots, n)$$

one which minimizes an integral

$$I = \int_{x_1}^{x_2} f(x, y, y')\,dx.$$

We can transform such a problem into a Lagrange problem by introducing new variables

(71) $$z_i(x) = \int_{x_1}^{x} g_i(x, y, y')\,dx.$$

The problem just stated is then equivalent to that of finding in the class of arcs

$$y = y(x), \quad z_i = z_i(x) \qquad (i = 1, \cdots, n;\ x_1 \leq x \leq x_2)$$

satisfying the conditions

(72) $$\begin{aligned} &g_i(x, y, y') - z_i' = 0, \\ &y(x_1) = y_1, \quad y(x_2) = y_2, \\ &z_i(x_1) = 0, \quad z_i(x_2) = l_i, \end{aligned} \qquad (i = 1, \cdots, n)$$

one which minimizes I.

The function F for a normal minimizing arc for this problem has the form

(73) $$F = f + \lambda_i(g_i - z_i')$$

and the differential equations determining such an arc are

(74) $$F_y - (d/dx)F_{y'} = 0$$

and the n equations

$$F_{z_i} - (d/dx) F_{z_i'} = (d\lambda_i/dx) = 0$$

which show that the multipliers λ_i are in this case all constants. The solutions of equations (74) form a family of the type

$$y = y(x, a, b, \lambda_1, \cdots, \lambda_n).$$

It contains $n+2$ arbitrary constants, and that is precisely the number of relations which the end-conditions (72) impose upon them as one readily verifies. It is evident that the equation (74) is unaltered if we think of the function F in it as defined by the equation

$$F = f + \lambda_i g_i \tag{75}$$

instead of equation (73).

For a minimizing arc which is not normal there would be a function F defined by equation (75) without the first term. It is clear that the equation (74) would then be defining the minimizing arcs for the problem of minimizing one of the integrals (70), say the first one, in the class of curves joining 1 with 2 and keeping the others constant. An arc E_{12} satisfying equations (74) and these conditions would in general be a minimizing arc for this problem, and it is evident that in that case there could be no other arc near E_{12} giving the first integral its minimum value l_1. Hence in a neighborhood of E_{12} the class of arcs joining 1 with 2 and satisfying conditions (70) would consist of E_{12} alone, and the original minimum problem would be a very trivial one in that neighborhood. Evidently the normal minimizing arcs are by far the most important ones. A similar but somewhat more complicated argument justifies the definition of normal minimizing arcs for the general Lagrange problem given in the preceding sections.

13. *The hanging chain.* It is a principle of mechanics that a chain suspended on two pegs will hang so that its center of gravity is as low as possible. In Section 2 it was seen that the form of the chain is therefore that of a minimizing arc for the problem in which we seek among the arcs $y = y(x)$ $(x_1 \leq x \leq x_2)$ satisfying the conditions

$$y(x_1) = y_1, \quad y(x_2) = y_2, \quad \int_{x_1}^{x_2} (1 + y'^2)^{1/2} dx = l, \tag{76}$$

one which minimizes the integral

$$I = \int_{x_1}^{x_2} y(1 + y'^2)^{1/2} dx.$$

The function F for a minimizing arc has the form

$$F = (y + \lambda)(1 + y'^2)^{1/2}$$

and since λ is now constant the differential equation (74) is equivalent to

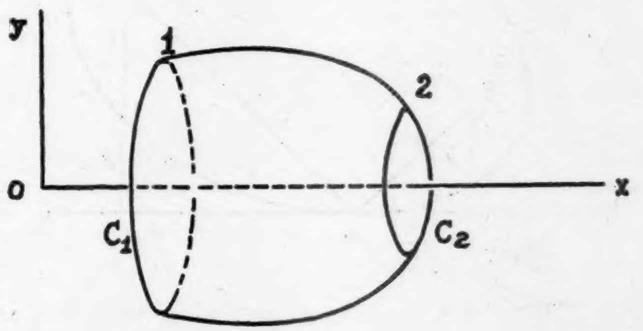
$$F - y'F_{y'} = (y+\lambda)/(1+y'^2)^{1/2} = b.$$

The integration of this equation has been many times discussed * and its solutions are the catenaries

$$y+\lambda = b\, ch[(x-a)/b].$$

This is a larger family than that of the catenaries for the problem of finding a minimum surface of revolution since it contains an arbitrary constant λ besides a and b. The extra constant is needed, however, for the problem of the hanging chain since there are three conditions (76) to be satisfied for that problem instead of the first two only.

14. *Soap films enclosing a given volume.* Let C_1 and C_2 be two circular discs with a common axis whose edges are joined by a soap film. It is well known that when the volume of air inclosed by the discs and the film is a



fixed constant k the form of the film surface will be that of a surface of revolution enclosing the volume k and having a minimum surface area. To determine the shape of the film we must seek therefore among the arcs $y = y(x)$ $(x_1 \leq x \leq x_2)$ satisfying the conditions

$$y(x_1) = y_1, \quad y(x_2) = y_2, \quad \int_{x_1}^{x_2} y^2 dx = k/\pi$$

one which minimizes the integral

$$I = \int_{x_1}^{x_2} y(1+y'^2)^{1/2} dx.$$

* See, for example, Bliss [5, p. 91].

The function F is $F = y(1+y'^2)^{1/2} + \lambda y^2$ and the equation (74) is equivalent to

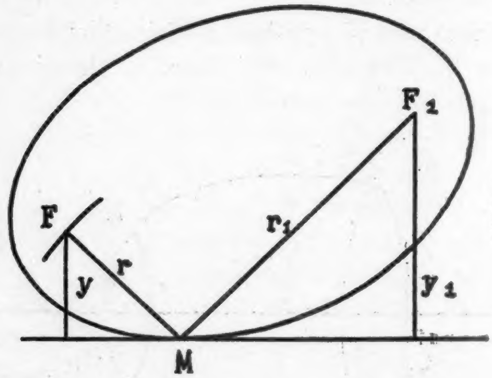
$$(77) \qquad F - y'F_{y'} = y/(1+y')^{1/2} - \lambda y^2 = c.$$

If we solve this equation for y' and separate the variables we find the solution in the form

$$x = \int \{(c - \lambda y^2)/[y^2 - (c - \lambda y^2)^2]^{1/2}\} dy + d.$$

The integral here is an elliptic integral which can be treated by well known methods.

The solutions of equations (77) can be characterized geometrically in an interesting fashion.* If an ellipse rolls on a straight line, as in the accom-



panying figure, its focus F describes a curve whose tangent is at every point perpendicular to FM. The coördinates (x, y) of F, and (x_1, y_1) of F_1, therefore satisfy the equations

$$y = r(dx/ds), \quad y_1 = r_1(dx/ds)$$

since by a well known property of the ellipse the angles made by r and r_1 with the tangent at M are equal. The equations

$$r + r_1 = 2a, \quad yy_1 = b^2$$

express two further well known properties of an ellipse, and elimination of r, r_1, y_1 from these and the preceding ones gives the differential equation

$$y^2 - 2ay(dx/ds) + b^2 = 0$$

* See, for example, Moigno-Lindelöf [6, p. 220].

for the locus of the point F. Equation (77) is identical with this if we set $\lambda = -1/2a$, $c = b^2/2a$. It can similarly be shown that for suitable determinations of λ and c equation (77) is also satisfied by the locus of the focus of a parabola or a hyperbola which rolls on the x-axis. The curves generated as described above by the foci of conics rolling on the x-axis are called unduloids and nodoids.

15. *The case when the functions ϕ_α contain no derivatives.* The problem of this section is that of finding among the arcs

$$y_i = y_i(x) \qquad (i = 1, \cdots, n;\ x_1 \leqq x \leqq x_2) \tag{78}$$

joining the two given points 1 and 2 and satisfying a set of equations of the form

$$\phi_\alpha(x, y_1, \cdots, y_n) = 0 \qquad (\alpha = 1, \cdots, m < n)$$

one which minimizes an integral

$$I = \int_{x_1}^{x_2} f(x, y, y')\, dx.$$

Let E_{12} be a particular arc whose minimizing properties are to be studied. It is always presupposed that in a neighborhood of the set of elements (x, y, y') on E_{12} the functions f, ϕ_α have continuous partial derivatives, say of the first four orders, and that the matrix $\|\partial\phi_\alpha/\partial y_i\|$ has rank m at every point of E_{12}.

In order to give this problem the usual Lagrange form we replace it by an equivalent one as follows. We may suppose without loss of generality that at the point 2 the determinant $|\partial\phi_\alpha/\partial y_\beta|$ is one of those of the matrix $\|\partial\phi_\alpha/\partial y_i\|$ which is different from zero. Then we seek to find among the arcs (78) satisfying the conditions

$$d\phi_\alpha/dx = \phi_{\alpha x} + \phi_{\alpha y_i} y_i' = 0, \tag{79}$$

$$x_1 - \alpha_1 = y_{i1} - \beta_{i1} = x_2 - \alpha_2 = y_{r2} - \beta_{r2} = 0 \tag{80}$$
$$(i = 1, \cdots, n;\ r = m + 1, \cdots, n)$$

one which minimizes I. The coördinates (α_1, β_{i1}) and (α_2, β_{i2}) are those of the points 1 and 2 and necessarily satisfy the equations $\phi_\alpha = 0$. The new problem is evidently equivalent to the old one, at least in a neighborhood of E_{12}, since every arc (78) which joins 1 with 2 and satisfies the equations $\phi_\alpha = 0$ also satisfies (79) and (80); and since, conversely, every arc sufficiently near E_{12} and satisfying (79) and (80) will also satisfy the equations

$\phi_\alpha = 0$ and pass through 1 and 2. This follows because the last $n - m + 1$ equations (80) and the equations $\phi_\alpha = 0$ at 2 imply $y_{\alpha 2} - \beta_{\alpha 2} = 0$.

Every extremal arc for the new problem is necessarily normal. The determinant analogous to (49) for the end-conditions (80) is in fact

$$\begin{vmatrix} \xi_{11} & \cdots & \xi_{1p} \\ \eta_{11}(x_1) & \cdots & \eta_{1p}(x_1) \\ \cdot & \cdots & \cdot \\ \eta_{n1}(x_1) & \cdots & \eta_{np}(x_1) \\ \xi_{21} & \cdots & \xi_{2p} \\ \eta_{m+1,1}(x_2) & \cdots & \eta_{m+1,p}(x_2) \\ \cdot & \cdots & \cdot \\ \eta_{n1}(x_2) & \cdots & \eta_{np}(x_2) \end{vmatrix}$$

where $p = 2n - m + 2$, and we can prove that the sets $\xi_{1\sigma}$, $\xi_{2\sigma}$, $\eta_{i\sigma}(x)$ $(\sigma = 1, \cdots, p)$ can be chosen so that this determinant is different from zero. The equations of variation are in fact readily seen to be the equations

$$(d/dx)\phi_{\alpha y_i}\eta_i = 0$$

which are equivalent to the system

$$\phi_{\alpha y_i}(x)\eta_i(x) = \phi_{\alpha y_i}(x_1)\eta_i(x_1). \tag{81}$$

If the end-values $\eta_i(x_1)$, $\eta_r(x_2)$ are selected arbitrarily these equations determine uniquely the end-values $\eta_\alpha(x_2)$ since the determinant $|\partial\phi_\alpha/\partial y_\beta|$ is by hypothesis different from zero at the point 2. Then the equations (81) and

$$\phi_{r y_i}(x)\eta_i(x) = \zeta_r(x), \tag{82}$$

where the auxiliary functions $\phi_r(x, y)$ are chosen so that the determinant $|\partial\phi_i/\partial y_k|$ is different from zero along E_{12}, determine the end-values $\zeta_r(x_1)$, $\zeta_r(x_2)$ uniquely when $\eta_i(x_1)$, $\eta_r(x_2)$ are given. If functions $\zeta_r(x)$ are chosen with the end-values $\zeta_r(x_1)$, $\zeta_r(x_2)$ but otherwise arbitrarily then equations (81) and (82) determine uniquely a corresponding set of variations $\eta_i(x)$ with the arbitrarily prescribed end-values $\eta_i(x_1)$, $\eta_r(x_2)$. Since ξ_1 and ξ_2 are arbitrary it is evident that the sets $\xi_{1\sigma}$, $\xi_{2\sigma}$, $\eta_{i\sigma}(x)$ can be chosen so that the determinant above is different from zero.

The function F for the Euler-Lagrange multiplier rule of the new problem can be taken in the form

$$F = f + \mu_\alpha(\phi_{\alpha x} + \phi_{\alpha y_k}y_k').$$

By a simple calculation the Euler-Lagrange equations are found to be

$$f_{y_i} - (d/dx)f_{y_i'} - \mu_\alpha'\phi_{\alpha y_i} = 0.$$

If we set $\lambda_a = -\mu_a'$ these are equivalent to the Euler-Lagrange equations calculated for the function

$$F = f + \lambda_a \phi_a$$

and we have the following result:

For the problem of finding among the arcs $y_i = y_i(x)$ $(i = 1, \cdots, n;$ $x_1 \leq x \leq x_2)$ *joining two given points and satisfying the equations*

$$\phi_a(x, y) = 0 \qquad (\alpha = 1, \cdots, m < n)$$

one which minimizes the integral

$$I = \int_{x_1}^{x_2} f(x, y, y')\, dx,$$

the extremal arcs all satisfy $n + m$ *equations of the form*

$$F_{y_i} - (d/dx) F_{y_i'} = 0, \quad \phi_a = 0$$

where F *is a function of the form* $F = f + \lambda_a \phi_a$.

16. *Geodesics on a surface.** The problem of finding the shortest curve joining two given points on a surface is analytically that of finding among the arcs

$$x = x(t), \quad y = y(t), \quad z = z(t) \qquad (t_1 \leq t \leq t_2)$$

satisfying the equation

$$\phi(x, y, z) = 0 \tag{83}$$

of the surface and joining the two given points, one which minimizes the integral

$$I = \int_{t_1}^{t_2} (x'^2 + y'^2 + z'^2)^{1/2} dt.$$

The function F for this problem, according to the results of the last section, is

$$F = (x'^2 + y'^2 + z'^2)^{1/2} + \lambda \phi$$

and the Euler-Lagrange equations are $\phi = 0$ and

$$\begin{aligned}
(d/dt)F_{x'} - F_x &= d/dt[x'/(x'^2 + y'^2 + z'^2)^{1/2}] - \lambda \phi_x = 0, \\
(d/dt)F_{y'} - F_y &= d/dt[y'/(x'^2 + y'^2 + z'^2)^{1/2}] - \lambda \phi_y = 0, \\
(d/dt)F_{z'} - F_z &= d/dt[z'/(x'^2 + y'^2 + z'^2)^{1/2}] - \lambda \phi_z = 0.
\end{aligned}$$

If these are written in the form

* Bolza [3, p. 553].

$$d^2x/ds^2 = \mu\phi_x, \quad d^2y/ds^2 = \mu\phi_y, \quad d^2z/ds^2 = \mu\phi_z, \quad \phi = 0,$$

where s is the length of arc, they express the fact that at each point of a minimizing arc the principal normal of the arc must coincide with the normal to the surface. Curves which have this property are called *geodesic lines* on the surface. Shortest arcs on a surface must always be sought among the geodesics.

For a sphere the equation (83) has the form

$$x^2 + y^2 + z^2 - 1 = 0$$

and the further equations of the geodesics are

$$d^2x/ds^2 = \mu x, \quad d^2y/ds^2 = \mu y, \quad d^2z/ds^2 = \mu z. \tag{84}$$

Let us determine constants a, b, c so that the expression

$$u = ax + by + cz$$

vanishes with its first derivative at one point of a geodesic on the sphere. Then u must be identically zero on the geodesic since the equation $u_{ss} = \mu u$ is a consequence of equations (84), and since the only solution of this last equation which can vanish with its derivative is $u \equiv 0$. It follows readily that the geodesics on a sphere are great circles cut out of the sphere by the planes $u = 0$.

17. *Brachistochrone on a surface.** Consider a particle of mass m moving in a field of force of such nature that when the particle is at the point (x, y, z) the force acting on it has the projections

$$mX = m(\partial U/\partial x), \quad mY = m(\partial U/\partial y), \quad mZ = m(\partial U/\partial z) \tag{85}$$

on the three coördinate axes, where U is a function of the coördinates x, y, z only. A constant gravitational field in the direction of the negative z-axis, for example, would have

$$X = 0, \quad Y = 0, \quad Z = -g, \quad U = -gz.$$

If a particle were constrained to move on a curve in such a field we should have the force in the direction of the tangent expressed in the two forms

$$mv' = m\,[X(dx/ds) + Y(dy/ds) + Z(dz/ds)]$$

where v is the velocity in the tangent direction, s is the length of arc measured along the curve, and the prime denotes a derivative with respect to the time t. Since $v = ds/dt$ this gives

* Moigno-Lindelöf [6, p. 301].

$$vv' = Xx' + Yy' + Zz' = U', \quad v^2 = 2U + c = 2(U - U_1) + v_1^2, \tag{86}$$

where U_1 and v_1 are values of U and v at an initial point 1. For a particle started at 1 with the velocity v_1 the velocity v at a point (x, y, z) is evidently a function of x, y, z and the same for all arcs joining 1 with this point. For an arc

$$x = x(t), \quad y = y(t), \quad z = z(t) \qquad (t_1 \leq t \leq t_2) \tag{87}$$

joining two fixed points 1 and 2 the time of descent of a particle starting at 1 with the velocity v_1 is

$$T = \int_{s_1}^{s_2} ds/v = \int_{t_1}^{t_2} (1/v)(x'^2 + y'^2 + z'^2)^{1/2} dt$$

where v is the function of x, y, z defined in equation (86).

The problem of finding an arc of quickest descent from a point 1 to a point 2 on a surface

$$\phi(x, y, z) = 0 \tag{88}$$

for a particle starting at 1 with a given velocity v_1 is equivalent analytically to that of finding among the arcs (87) joining the two given points and satisfying the equation (88), one which minimizes the integral T.

The function F for this problem is

$$F = (1/v)(x'^2 + y'^2 + z'^2)^{1/2} + \lambda\phi$$

and the Euler-Lagrange equations have the form

$$\begin{aligned}
\frac{d}{dt} F_{x'} - F_x &= \frac{d}{dt} \frac{1}{v} \frac{dx}{ds} + \frac{v_x}{v^2} \frac{ds}{dt} - \lambda\phi_x = 0, \\
\frac{d}{dt} F_{y'} - F_y &= \frac{d}{dt} \frac{1}{v} \frac{dy}{ds} + \frac{v_y}{v^2} \frac{ds}{dt} - \lambda\phi_y = 0, \\
\frac{d}{dt} F_{z'} - F_z &= \frac{d}{dt} \frac{1}{v} \frac{dz}{ds} + \frac{v_z}{v^2} \frac{ds}{dt} - \lambda\phi_z = 0
\end{aligned}$$

to which must be adjoined the equation $\phi = 0$. When multiplied through by dt/ds the equations above become

$$\begin{aligned}
-(v_s/v^2)x_s + (1/v)x_{ss} + (v_x/v^2) - \mu\phi_x &= 0, \\
-(v_s/v^2)y_s + (1/v)y_{ss} + (v_y/v^2) - \mu\phi_y &= 0, \\
-(v_s/v^2)z_s + (1/x)z_{ss} + (v_z/v^2) - \mu\phi_z &= 0.
\end{aligned}$$

Multiplied respectively by the direction cosines l, m, n of the direction tan-

gent to the surface, perpendicular to the extremal, and making an acute angle with its principal normal, these give

$$(1/v)(lx_{ss}+my_{ss}+nz_{ss})+(1/v^2)(v_x l+v_y m+v_z n)=0$$

from which we can show that

$$(v^2/\rho)\cos\alpha+R\cos\beta=0 \tag{89}$$

where ρ is the radius of curvature of the curve, α the angle between the radius and the direction $l:m:n$, R the total impressed force, and β the angle between the force and $l:m:n$. This result follows immediately since the numbers ρx_{ss}, ρy_{ss}, ρz_{ss} are the three direction cosines of the principal normal to the curve on which the radius ρ lies, and since from equations (86)

$$vv_x=U_x,\quad vv_y=U_y,\quad vv_z=U_z$$

and U_x, U_y, U_z are the projections on the coördinate axes of the force R. The equation (89) justifies the following characteristic property of brachistochrones on a surface:

Consider a surface $\phi(x,y,z)=0$ lying in a field of force whose vector at (x,y,z) has magnitude R and components X, Y, Z defined by a force function $U(x,y,z)$, as indicated in equations (85). *The centrifugal force of a particle moving on a curve is by definition directed in the direction opposite to that of the radius ρ of the first curvature, and has magnitude v^2/ρ where v is the velocity of the particle. Equation* (89) *shows that at each point of a brachistochrone curve on the surface $\phi=0$ the projection of the centrifugal force on the particular normal to the curve which is also tangent to the surface, is equal to the projection on that same line of the impressed force R.*

This is a characteristic property of brachistochrones. Equation (89) shows that the radius of geodesic curvature $\rho_g=\rho\sec\alpha$ is defined by the equation

$$1/\rho_g=-(R/v^2)\cos\beta. \tag{90}$$

On a surface whose equations are in parametric form with parameters u, v the geodesic curvature of an arc defined by an equation $v=v(u)$ is expressed in terms of $v(u)$, $v'(u)$, $v''(u)$ while the quantities in the second members of the last equation involve only $v(u)$ and $v'(u)$. This equation is consequently a differential equation of the second order. Through each point and direction on the surface there passes therefore one and only one extremal arc for the brachistochrone problem. One can readily verify that the equation

(90) is satisfied by the brachistochrones on a plane which are the well-known cycloids.

18. *The curve of equilibrium of a chain hanging on a surface.** Let us accept from the theories of mechanics the statement that the potential energy of a chain of the form

$$x = x(t), \quad y = y(t), \quad z = z(t) \qquad (t_1 \leq t \leq t_2) \tag{91}$$

in a field of force like the one described in the last section is

$$P = -\int_{s_1}^{s_2} U ds = -\int_{t_1}^{t_2} U(x'^2 + y'^2 + z'^2)^{1/2} dt,$$

and the statement that a chain at rest will be in equilibrium when the potential energy is a minimum. The problem of finding the position of equilibrium of a chain of given length l joining two given points 1 and 2 and lying on a surface

$$\phi(x, y, z) = 0$$

in such a field is then that of finding among the arcs (91) joining 1 with 2 and satisfying the conditions

$$\int_{t_1}^{t_2} (x'^2 + y'^2 + z'^2)^{1/2} dt = l, \quad \phi(x, y, z) = 0$$

one which minimizes the integral P. In a gravitational field the value of U is $-gz$.

This problem is partly of the isoperimetric and partly of the Lagrange type. By methods used above one readily verifies that its function F now has the form

$$F = (U + \lambda)(x'^2 + y'^2 + z'^2)^{1/2} + \mu\phi,$$

where λ is a constant, and that its extremal arcs satisfy $\phi = 0$ and the equations

$$\begin{aligned}
d/dt[(U + \lambda)x'/(x'^2 + y'^2 + z'^2)^{1/2}] - U_x(x'^2 + y'^2 + z'^2)^{1/2} - \mu\phi &= 0,\\
d/dt[(U + \lambda)y'/(x'^2 + y'^2 + z'^2)^{1/2}] - U_y(x'^2 + y'^2 + z'^2)^{1/2} - \mu\phi_y &= 0,\\
d/dt[(U + \lambda)z'/(x'^2 + y'^2 + z'^2)^{1/2}] - U_z(x'^2 + y'^2 + z'^2)^{1/2} - \mu\phi_z &= 0.
\end{aligned}$$

These are equivalent to

$$\begin{aligned}
U_s x_s + (U + \lambda)x_{ss} - U_x - \nu\phi_x &= 0,\\
U_s y_s + (U + \lambda)y_{ss} - U_y - \nu\phi_y &= 0,\\
U_s z_s + (U + \lambda)z_{ss} - U_z - \nu\phi_z &= 0.
\end{aligned}$$

---

* Moigno-Lindelöf [6, p. 313].

Multiplied respectively by the direction cosines l, m, n of the direction tangent to the surface, perpendicular to the extremal, and making an acute angle with its principal normal, these give

$$(U+\lambda)\cos\alpha/\rho = R\cos\beta,$$

or

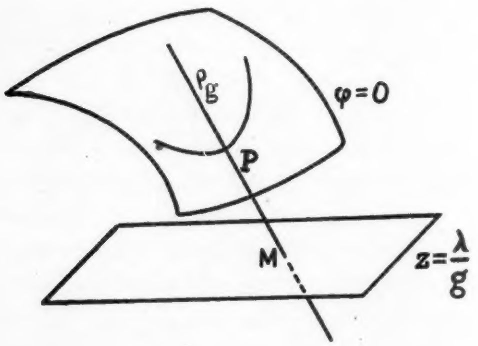
$$\rho_g = (U+\lambda)\sec\beta/R,$$

where ρ, ρ_g, α, β have the significance of the last section. Like the equation (90) this defines a two-parameter family of extremals arcs on the surface $\phi = 0$.

For the particular case of a gravitational field of force $U = -gz$, $R = g$, and β is the angle between the negative z-axis and the direction $l:m:n$ so that $\cos\beta = -n$. Hence in this case

$$\rho_g = [(z-\lambda)/g]/n$$

which says that at each point of a curve of equilibrium the radius of geodesic curvature is equal to the segment PM in the figure, bounded on the line



$l:m:n$ perpendicular to the curve and tangent to the surface $\phi = 0$ by the point P and the plane $z = \lambda/g$. This is a well known property of a catenary $y = c + b\, ch\, [(x-a)/b]$, which is the curve of a hanging chain in a vertical plane. The surface $\phi = 0$ is in this case the xy-plane, the radius ρ_g is the radius of curvature of the catenary, and the plane $z = \lambda/g$ is to be represented by the line $y = c$. The radius of curvature at a point P of the catenary is equal to the intercept on the normal to the catenary at P between the point P and the line $y = c$.

19. *Hamilton's principle.** Suppose that the n particles whose coördinates and masses are x_i, y_i, z_i, m_i $(i = 1, \cdots, n)$ move in a field of force

---

* Bolza [3, p. 554].

in space such that the force acting at any instant on the i-th particle has components

$$X_i = U_{x_i}, \quad Y_i = U_{y_i}, \quad Z_i = U_{z_i},$$

where U is a function of the time t and the $3n$ coördinates x_i, y_i, z_i. Suppose further that the motions of the particles are restricted by conditions of the form

$$\phi_a = 0 \qquad (\alpha = 1, \cdots, m < 3n),$$

where the functions ϕ_a also depend upon t and the coördinates. The differential equations of motion of the particles, as established in treatises in mechanics, are

$$\begin{aligned} m_i x_i'' &= U_{x_i} + \Sigma_a \lambda_a \phi_{a x_i}, \\ m_i y_i'' &= U_{y_i} + \Sigma_a \lambda_a \phi_{a y_i}, \\ m_i z_i'' &= U_{z_i} + \Sigma_a \lambda_a \phi_{a z_i}, \end{aligned} \tag{92}$$

where α has the range from 1 to m. In this and the following sections of this chapter sums will be indicated as usual and no umbral indices will be used.

Hamilton's principle is simply the statement that the differential equations (92) are the differential equations of the minimizing arcs of the problem of finding in the class of $3n$-dimensional arcs

$$x_i = x_i(t), \quad y_i = y_i(t), \quad z_i = z_i(t) \qquad (t_1 \leq t \leq t_2;\ i = 1, \cdots, n)$$

joining two given points and satisfying the equations $\phi_a = 0$, one which minimizes the integral

$$I = \int_{t_1}^{t_2} (T + U)\, dt$$

where U is the force function and T the so-called kinetic energy

$$T = \tfrac{1}{2}\Sigma_i m_i v_i^2 = \tfrac{1}{2}\Sigma_i m_i (x_i'^2 + y_i'^2 + z_i'^2).$$

It is very easy to show that the equations (92) are the Euler-Lagrange equations for this problem. We have only to set up these equations for the function

$$F = T + U + \Sigma_a \lambda_a \phi_a.$$

An important application of Hamilton's principle is that of determining the equations of motion in terms of the so-called generalized coördinates of Lagrange. The number of coördinates x_i, y_i, z_i is $3n$ and the number of equations $\phi_a = 0$ is m. It is in general possible in an infinity of ways to express these coördinates as functions of t and $3n - m$ arbitrary parameters $q_1, \cdots, q_{3n-m}$ satisfying identically the equations $\phi_a = 0$ and giving all the solutions of these equations. The functions T and U then take the form

$$T = T(t, q, q'), \quad U = U(t, q),$$

and the problem is transformed into that of finding among the arcs $q_r = q_r(t)$ $(r = 1, \cdots, 3n - m)$ joining the two given points one which minimizes the integral I. No adjoined conditions $\phi_a = 0$ are now necessary. The differential equations of the minimizing arcs for the new problem are the equations

$$\frac{d}{dt}\frac{\partial T}{\partial q_r'} - \frac{\partial}{\partial q_r}(T + U) = 0 \qquad (r = 1, \cdots, 3n - m).$$

The important result is that the form of these equations is the same no matter what new coördinates $q_1, \cdots, q_{3n-m}$ with the properties described above are used.

20. *Two forms of the principle of least action.** Let us now consider the somewhat special case where the functions U and ϕ_a of the last section do not contain the time t explicitly. If the equations (92) are multiplied by x_i', y_i', z_i', respectively, added, and integrated we find the well-known relation

$$T = U + h$$

where h is a constant of integration. This is the principle of the conservation of energy which says that the sum of the kinetic energy T and the potential energy $-U$ of a system satisfying equations (92) is always a constant.

Jacobi's form of the principle of least action states that the totality of dynamical trajectories satisfying equations (92) and having a given energy constant h is identical with the totality of extremals for the problem of finding among the arcs

$$x_i = x_i(u), \quad y_i = y_i(u), \quad z_i = z_i(u) \qquad (i = 1, \cdots, n;\ u_1 \leq u \leq u_2)$$

joining two given points and satisfying the equations $\phi_a = 0$ one which minimizes the integral

$$I = \int_{u_1}^{u_2} [2(U + h)S]^{1/2} du,$$

where S is simply a notation for the sum

$$S = \Sigma_i m_i (x_{iu}^2 + y_{iu}^2 + z_{iu}^2).$$

The parameter u is not in this case the time, but if at the time t_0 the particles are at the places defined on their trajectories by the parameter value u_0, then it turns out that the time at the place defined by u is

* Bolza [3, pp. 556, 586].

$$t = t_0 + \int_{u_0}^{u} \{S/[2(U+h)]\}^{1/2} du, \tag{93}$$

as one would expect from the relation $S(du/dt)^2 = 2T = 2(U+h)$.

To prove these statements we note that the function F for the minimizing problem just described is

$$F = [2(U+h)S]^{1/2} + \Sigma_a \mu_a \phi_a.$$

A typical one of the Euler-Lagrange equations is

$$(d/du)\{[2(U+h)]/S\}^{1/2} m_i x_{iu} - U_{x_i}\{S/[2(U+h)]\}^{1/2} - \Sigma_a \mu_a \phi_{ax_i} = 0.$$

If we introduce the parameter t along a solution of this equation by means of the formula (93) then the equation itself takes the form

$$m_i x_i'' - U_{x_i} - \Sigma_a \lambda_a \phi_{ax_i} = 0$$

when $\lambda_a = \mu_a\ (du/dt)$, which is the same as the first equation (92).

Lagrange's form of the principle of least action is again a principle for describing those mechanical trajectories which satisfy equations (92) and have a given energy constant h. They are extremals for the problem of finding among the arcs

$$x_i = x_i(t), \quad y_i = y_i(t), \quad z_i = z_i(t) \qquad (i = 1, \cdots, n;\ t_1 \leq t \leq t_2)$$

passing through given initial values of the coördinates for a given initial time t_1, passing through given end-values of the coördinates for an unspecified time t_2, and satisfying the equations

$$T - U - h = 0, \quad \phi_a = 0 \tag{94}$$

one which minimizes the integral

$$I = \int_{t_1}^{t_2} T dt.$$

This is a problem with a variable second end-point since t_2 is not specified. The function F for it is

$$F = T + \lambda(T - U - h) + \Sigma_a \mu_a \phi_a$$

and a typical Lagrange equation is

$$(d/dt)(1+\lambda) m_i x_i' + \lambda U_{x_i} - \Sigma_a \mu_a \phi_{ax_i} = 0. \tag{95}$$

When this equation is multiplied by x_i' and added to the other similar ones, it is found with the help of equations (94) that

$$\lambda = (k/2T) - 1/2$$

where k is a constant.

If all the end-values except x_2 are fixed in the theorem of pages 692-3, then the matrix (48) is square and its vanishing requires that

$$F(x_2) - \Sigma_i y_{i2}' F_{y_i'}(x_2) = 0.$$

Interpreted for the function F above this gives $\lambda = -1/2$ at $t = t_2$, with the help of equations (94). It follows that in the formula deduced above for λ we must have $k = 0$ and hence that $\lambda = -1/2$ for all values of t. Equation (95) then takes the form of the first equation (92) when we set $\lambda_a = 2\mu_a$.

CHAPTER III.

FURTHER NECESSARY CONDITIONS FOR A MINIMUM.

In this third chapter three further necessary conditions on a minimizing arc for the Lagrange problem will be developed, analogous to those of Weierstrass, Legendre, and Jacobi for the simpler types of problems of the calculus of variations. The analogue of Legendre's condition was first deduced by Clebsch [20] and the analogue of Jacobi's condition by A. Mayer[24]. For the deduction of these necessary conditions and for a number of other purposes we shall find the auxiliary theorems of the next section convenient.

21. *Two important auxiliary theorems.* Consider a one parameter family of admissible arcs

$$(96) \qquad y_i = y_i(x, b), \quad x_3(b) \leq x \leq x_4(b), \qquad (i = 1, \cdots, n)$$

for which the functions $x_3(b)$, $x_4(b)$, $y_i(x, b)$, $y_i'(x, b)$ are continuous and have continuous derivatives with respect to b in the domain of values (x, b) defined by the inequalities $b' \leq b \leq b''$, $x_3(b) \leq x \leq x_4(b)$, and whose end values describe two arcs C and D. The values of I taken along the arcs (96) are given by the formula

$$I(b) = \int_{x_3}^{x_4} f[x, y(x, b), y'(x, b)]\,dx$$

which has the derivative

$$I'(b) = f x_b \Big|_3^4 + \int_{x_3}^{x_4} \{f_{y_i} y_{ib} + F_{y_i'} y'_{ib}\}\,dx.$$

The index here is umbral and we shall use umbral indices freely elsewhere in this chapter. Since the arcs (96) are all admissible this result may also be written in the form

$$\lambda_0 I'(b) = F x_b \Big|_3^4 + \int_{x_3}^{x_4} \{F_{y_i} y_{ib} + F_{y_i'} y'_{ib}\} dx, \tag{97}$$

where the multipliers λ_0, $\lambda_a(x)$ in the function

$$F = \lambda_0 f + \lambda_a \phi_a$$

are entirely arbitrary. If now a particular arc of the family (96) satisfies the equations

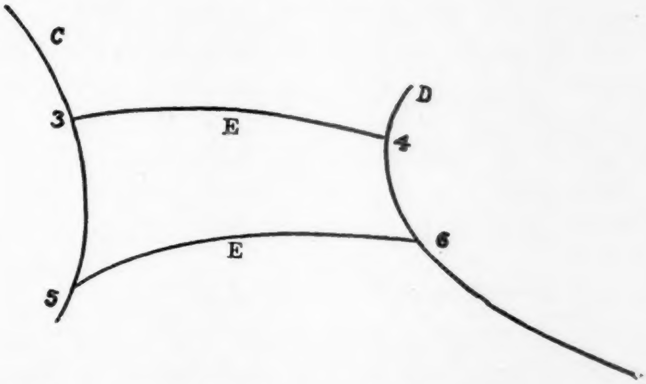
$$F_{y_i'} = \int_{x_1}^{x} F_{y_i} dx + c_i$$

with a set of multipliers λ_0, $\lambda_a(x)$, then the introduction of these multipliers enables us to replace formula (97) by

$$\lambda_0 I'(b) = F x_b + F_{y_i'} y_{ib} \Big|_3^4 \quad .$$

where b is the particular value defining that arc. Since the equations of C and D are deduced from

$$x = x(b), \quad y_i = y_i[x(b), b]$$



by replacing $x(b)$ by $x_3(b)$ and $x_4(b)$, respectively, it follows that along either of these arcs

$$dy_i = y'_i dx + y_{ib} db,$$

and therefore that

$$\lambda_0 dI = F dx + (dy_i - y_i' dx) F_{y_i'} \Big|_3^4.$$

Hence we have the following theorem:

AUXILIARY THEOREM. I. *Let*

$$y_i = y_i(x, b), \quad x_1(b) \leq x \leq x_2(b), \qquad (i = 1, \cdots, n) \tag{98}$$

be a one-parameter family of admissible arcs without corners whose end-points describe two arcs C and D. If one of the arcs (98) *satisfies the equations*

$$F_{y_i'} = \int_{x_1}^{x} F_{y_i} dx + c_i \tag{99}$$

with a set of multipliers λ_0, $\lambda_\alpha(x)$ then for the value of b defining it the values of I along the arcs (98) *have a differential defined by the equation*

$$\lambda_0 dI = F dx + (dy_i - y_i' dx) F_{y_i'} \Big|_3^4. \tag{100}$$

In this formula the differentials dx, dy_i at the point 3 are those of C, and at the point 4 those of D.

If the particular arc along which the equation (99) holds is a normal arc then λ_0 can be taken equal to unity in formula (100). If each of the curves (98) has a set of multipliers $\lambda_0(b)$, $\lambda_\alpha(x, b)$ with which it satisfies equations (99), then the formula (100) holds along every arc of the family. We suppose that the functions $\lambda_0(b)$, $\lambda_\alpha(x, b)$ are continuous for $b' \leq b \leq b''$, $x_3(b) \leq x \leq x_4(b)$, and then we have

AUXILIARY THEOREM II. *Suppose that the arcs of the family* (98) *are all extremal arcs with multipliers of the form $\lambda_0 = 1$, $\lambda_\alpha(x, b)$. Then the values of I on two arcs E_{34} and E_{56} of the family satisfy the equation*

$$I(E_{56}) - I(E_{34}) = I^*(D_{46}) - I^*(C_{35})$$

with the values of the integral

$$I^* = \int \{F dx + (dy_i - y_i' dx) F_{y'_i}\}$$

along the corresponding segments C_{35} and D_{46} shown in the last figure.

This is readily found by integrating both sides of formula (100) with respect to b from the value b' defining simultaneously the points 3 and 4 to the value b'' defining similarly 5 and 6. The integrand of the integral I^* is readily seen to be a continuous function of b on the arcs C_{35} and D_{46} corresponding to the interval $b'b''$, on account of the properties of the functions $x(b)$, $y_i(x, b)$ defining the family (98).

22. *Necessary conditions analogous to those of Weierstrass and Legendre.* Suppose that the equations

$$y_i = y_i(x) \qquad (i = 1, \cdots, n;\ x_1 \leq x \leq x_2)$$

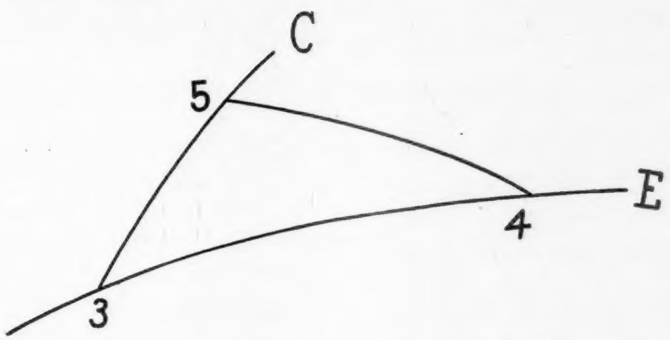
are those of a minimizing arc E_{12} for our problem.

We shall designate a set of values (x, y, y') as admissible if it lies in the neighborhood $\mathfrak{R}$ of page 1, satisfies the equations $\phi_\alpha = 0$, and gives the matrix

$\| \phi_{\alpha y_i'} \|$ the rank m. Let 3 be an arbitrary point on the arc E_{12} and let (x_3, y_{i3}, Y'_{i3}) be an admissible set. There is always an admissible arc

$$\text{(C)} \qquad y_i = Y_i(x) \qquad (x_3 \leq x \leq x_3 + h)$$

through this set since the equations $\phi_\alpha = 0$ determine uniquely m of the functions $Y_i(x)$ passing with their derivatives through the values prescribed by this initial set when the $n - m$ other functions $Y_i(x)$ have been chosen with initial values of themselves and their derivatives through their corresponding initial values of the set.



Suppose now that the arc E_{12} is normal on every sub-interval, and let 4 be so near to 3 on E_{12} that the arc E_{34} contains no corner. There is a $2n$-parameter family of admissible arcs $y_i = y_i(x, b_1, \cdots, b_{2n})$ containing E_{12} for $(b_1, \cdots, b_{2n}) = (0, \cdots, 0)$ and having $2n$ sets of variations $\eta_i(x)$ for which the determinant (38) with x_1, x_2 replaced by x_3, x_4 is different from zero. The $2n$ equations

$$y_i(x_5, b_1, \cdots, b_{2n}) = Y_i(x_5), \qquad y_i(x_4, b_1, \cdots, b_{2n}) = y_{i4}$$

have the initial solution $(x_5, b_1, \cdots, b_{2n}) = (x_3, 0, \cdots, 0)$ at which their functional determinant for $b_1, \cdots, b_{2n}$ is the determinant (38) with $x_1 = x_3$, $x_2 = x_4$ and different from zero. Hence they determine $2n$ functions $b_\mu = B_\mu(x_5)$ which vanish for $x_5 = x_3$. The family

$$y_i = y_i[x, B_1(x_5), \cdots, B_{2n}(x_5)] = y_i(x, x_5)$$

is now a one-parameter family of arcs joining the curve C of the figure to the point 4. The sum

$$\begin{aligned} \Phi(x_5) &= I(C_{35}) + I(E_{54}) \\ &= \int_{x_3}^{x_5} f(x, Y, Y')\,dx + \int_{x_5}^{x_4} f[x, y(x, x_5), y'(x, x_5)]\,dx \end{aligned}$$

must have its derivative ≥ 0 at x_3 if $I(E_{12})$ is to be a minimum. But with the help of formula (100) this derivative is seen to be

$$\Phi'(x_3) = E(x, y, y', Y', \lambda)\,|^3$$

if we define the E-function by the formula

$$(101)\quad E = F(x, y, Y', \lambda) - F(x, y, y', \lambda) - (Y_i' - y_i')F_{y_i'}(x, y, y', \lambda).$$

The multipliers in F are those associated uniquely with the normal minimizing arc E_{12}. Evidently one may always replace f by F for admissible sets (x, y, y'). We have then the following necessary condition:

ANALOGUE OF WEIERSTRASS NECESSARY CONDITION. *At each element (x, y, y', λ) of a minimizing arc which is normal on every sub-interval the inequality*

$$E(x, y, y', Y', \lambda) \geqq 0$$

must be satisfied for every admissible set $(x, y, Y') \neq (x, y, y')$.

The proof just given does not apply to the values x, y, y', λ at the right-hand end of an arc abutting on a corner, but it can be modified easily to be applicable by taking the point 4 at the left of 3, or one can infer the desired result by continuity considerations.

Consider now a set of values π_i satisfying the equations

$$(102)\qquad \phi_{\alpha y_i'}\pi_i = 0$$

at an element (x, y, y') of E_{12}. By means of the equations

$$(103)\qquad \phi_{r y_i'}\pi_i = \kappa_r$$

these define $n - m$ further quantities κ_r. The equations

$$\phi_\alpha(x, y, p) = 0, \qquad \phi_r(x, y, p) = z_r + \epsilon\kappa_r$$

now have the initial solution $(\epsilon, p_1, \cdots, p_n) = (0, y_1', \cdots, y_n')$ and determine uniquely a set of solutions $p_i(\epsilon)$ with initial values $p_i(0) = y_i'$. The derivatives $p_i'(0)$ of these functions satisfy equations (102) and (103) when inserted in place of the numbers π_i and hence must coincide with them. The sets $(x, y, p(\epsilon))$ are now all admissible for sufficiently small values of ϵ, and according to the last theorem must satisfy the condition

$$E(x, y, y', p(\epsilon), \lambda) \geqq 0.$$

But we readily verify that this expression vanishes with its first derivative for ϵ at the value $\epsilon = 0$. Its second derivative

$$F_{y_i' y_k'}\pi_i\pi_k$$

at $\epsilon = 0$ must therefore be $\geqq 0$, from which we infer the

NECESSARY CONDITION OF CLEBSCH. *At every element* (x, y, y', λ) *of a minimizing arc which is normal on every sub-interval the inequality*

$$F_{y_i' y_k'}(x, y, y', \lambda)\pi_i\pi_k \geq 0$$

must be satisfied by every set $(\pi_i, \cdots, \pi_n) \neq (0, \cdots, 0)$ *which is a solution of the m equations*

$$\phi_{\alpha y_i'}(x, y, y')\pi_i = 0.$$

23. *The envelope theorem.* According to the theorem of page 687, every extremal arc E_{12} along which the determinant R is different from zero is a member of a $2n$-parameter family of extremals of the form

$$y_i = y_i(x, a, b), \qquad \lambda_\alpha = \lambda_\alpha(x, a, b)$$

for special values a_{i0}, b_{i0} of the parameters. The family can be so chosen that the determinant (36) is different from zero at x_1, and we shall see in Section 27, page 727, that this determinant is in fact different from zero everywhere on E_{12}. If the constants a_i, b_i are replaced by functions $a_i(t)$, $b_i(t)$ with the initial values $a_i(0) = a_{i0}$, $b_i(0) = b_{i0}$ a one-parameter family of extremals is defined containing the arc E_{12} for the special parameter value $t = 0$. The arcs of this family will pass through the point 1 for $x = x_1$, and will touch an enveloping curve D at the points defined by a suitably chosen function $x(t)$, if the equations

$$\begin{aligned} x' = k, \quad & y_{ix}x' + y_{ia_k}a_k' + y_{ib_k}b_k' = ky_{ix}, \\ y_{i1} = {} & y_i(x_1, a, b) \end{aligned}$$

hold identically in t when x, a_i, b_i are replaced by the functions of t described above and the primes denote derivatives with respect to t. The first row of equations imposes the condition that the direction of the tangent to the curve D shall coincide with the direction $1 : y_1' : \cdots : y_n'$ of the tangent to the extremal. In order that these equations may be true it is evidently necessary and sufficient that the equations

$$\begin{aligned} y_{ia_k}[x(t), a(t), b(t)]\, a_k' + y_{ib_k}[x(t), a(t), b(t)]\, b_k' = 0, \\ y_{ia_k}[x_1, a(t), b(t)]\, a_k' + y_{ib_k}[x_1, a(t), b(t)]\, b_k' = 0, \end{aligned}$$

hold identically in t. If the derivatives a_k', b_k' are not zero it follows that the determinant

$$\Delta(x, x_1, a, b) = \begin{vmatrix} y_{ia_k}(x, a, b) & y_{ib_k}(x, a, b) \\ y_{ia_k}(x_1, a, b) & y_{ib_k}(x_1, a, b) \end{vmatrix} \tag{104}$$

vanishes identically in t when $x(t)$, $a_i(t)$, $b_i(t)$ are substituted.

DEFINITION OF A CONJUGATE POINT. A value $x_3 \neq x_1$ is said to define a *point 3 conjugate to 1 on the extremal arc* E_{12} if it is a root of a determinant $\Delta(x, x_1, a_0, b_0)$ belonging to a $2n$-parameter family of extremals $y_i = y_i(x, a, b)$, $\lambda_a = \lambda_a(x, a, b)$ for which the determinant

$$\begin{vmatrix} y_{ia_k} & y_{ib_k} \\ v_{ia_k} & v_{ib_k} \end{vmatrix}$$

is different from zero on E_{12} as described on page 727.

Suppose now that 3 is such a conjugate point, and furthermore one at which the derivative Δ_x does not vanish. It is evident that if $\Delta_x \neq 0$ one at least of the minors of order $2n - 1$ of Δ does not vanish at 3, and that the same property is therefore possessed by one at least of the determinants of order $2n$ of the matrix

$$\left\| \begin{matrix} \Delta_x & \Delta_{a_k} & \Delta_{b_k} \\ 0 & y_{ia_k}(x, a, b) & y_{ib_k}(x, a, b) \\ 0 & y_{ia_k}(x_1, a, b) & y_{ib_k}(x_1, a, b) \end{matrix} \right\|$$

since one at least of these determinants is the product of Δ_x by a non-vanishing minor of Δ. Then the first of the differential equations

$$\begin{aligned} \Delta_x(x, x_1, a, b)\,dx + \Delta_{a_k}(x, x_1, a, b)\,da_k + \Delta_{b_k}(x, x_1, a, b)\,db_k &= 0, \\ y_{ia_k}(x, a, b)\,da_k + \quad y_{ib_k}(x, a, b)\,db_k &= 0, \\ y_{ia_k}(x_1, a, b)\,da_k + \quad y_{ib_k}(x_1, a, b)\,db_k &= 0, \end{aligned} \tag{105}$$

with $2n - 1$ of the others determine functions $x(t)$, $a_k(t)$, $b_k(t)$ with the initial values $x(0) = x_3$, $a_k(0) = a_{k0}$, $b_k(0) = b_{k0}$, and with derivatives x', a_k', b_k' not all zero at $t = 0$. Since $\Delta_x \neq 0$ at 3 it follows further that a_k', b_k' can not all vanish at $t = 0$. Since Δ vanishes at these initial values and has its derivative with respect to t identically zero, it must be itself identically zero in t. One sees readily then that the one remaining equation (105) is a consequence of the others when $x(t)$, $a_k(t)$, $b_k(t)$ are substituted. The following theorem is established:

Let E_{12} be an extremal arc along which the determinant R is different from zero, and let 3 be a point conjugate to 1 on E_{12} at which the derivative Δ_x of the determinant (104) *is different from zero. Then there exists through the point 1 a one-parameter family of extremals*

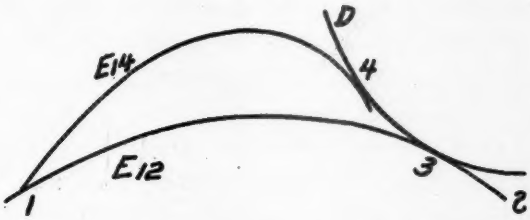
$$y_i = y_i(x, t), \quad \lambda_a = \lambda_a(x, t) \tag{106}$$

containing E_{12} for the parameter value $t = 0$ and having an envelope D which touches E_{12} at the point 3. The functions y_i, y_{ix}, λ_a and the function $x(t)$

defining D have continuous derivatives in a neighborhood of the values x, t belonging to the arc E_{12}.

The last statement of the theorem is a consequence of the hypothesis (b) of page 676. For as a result of this hypothesis the functions y_i, y_{ix}, λ_a of the theorem on page 687 have continuous derivatives of the second order at least, and the solutions $x(t)$, $a_k(t)$, $b_k(t)$ of the equations (105) must therefore have continuous derivatives of at least the first order.

THE ENVELOPE THEOREM. *If the envelope D of the one-parameter family of extremals* (106) *has a branch projecting backward from 3 toward the*



point 1, as shown in the figure, then for every position of the point 4 on D preceding and near to 3 the arc $E_{14} + D_{43} + E_{32}$ is an admissible arc satisfying the equations $\phi_a = 0$. Furthermore for every such arc

$$I(E_{14} + D_{43} + E_{32}) = I(E_{12}).$$

Expressed in integral form the value of $I(E_{14} + D_{43})$ is

$$I(E_{14} + D_{43}) = \int_{x_1}^{x(t)} f\,[x, y(x, t) y'(x, t)]\, dx + \int_t^0 f\, x'\, dt$$

where the arguments in f in the last integral are $x(t), y[x(t), t], y'[x(t), t]$. The differential of the first integral with respect to t is given by formula (100) of page 716, and that of the second integral is readily found. It follows that

$$dI(E_{14} + D_{43}) = -E(x, y, y', Y', \lambda)\, dx\, |^4$$

where Y' is the slope of D. But this vanishes identically in t since $Y' = y'$ at every point of D, and the final conclusion of the theorem is established. Evidently the envelope D satisfies the equations $\phi_a = 0$ at each point 4 since it is tangent at that point to the extremal arc E_{14}.

24. *The analogue of Jacobi's condition.* The analogue of Jacobi's condition was discovered for the Lagrange problem by A. Mayer. Its statement is as follows:

4

THE NECESSARY CONDITION OF MAYER. *Let E_{12} be an extremal arc for the Lagrange problem which is normal on every sub-interval of x_1x_2 and has the determinant*

$$R = \begin{vmatrix} F_{y_i'y_k'} & \phi_{\alpha y_i'} \\ \phi_{\alpha y_k'} & 0 \end{vmatrix}$$

different from zero at every point of it. If E_{12} is a minimizing arc for the problem then between 1 and 2 on E_{12} there can be no point 3 conjugate to 1.

The proof of the statement for the case when the envelope has a branch as described in the envelope theorem is not difficult if one accepts the assertion that every extremal arc of a family $y_i(x, a, b)$ whose end-values x_1, x_2 and parameters a, b are sufficiently near to those of a normal extremal arc of the family is also normal. The proof of this assertion depends upon the fact that when the functions $y_i(x, a, b)$ are substituted in the equations of variation, the solutions $\eta_i(x, a, b)$ of those equations are continuous in the parameters a, b as well as x. Hence if there are $2n$ sets of variations η_{is} $(s = 1, \cdots, 2n)$ making the determinant (38) different from zero for the values x_{10}, x_{20}, a_0, b_0 defining the normal extremal, then this determinant will remain different from zero for neighboring values x_1, x_2, a, b.

If the arcs $E_{14} + D_{43} + E_{32}$ of the envelope theorem were all minimizing arcs they would necessarily have continuous multipliers since they have no corners. According to the assertion discussed in the last paragraph those sufficiently near to E_{12} would be normal on the intervals x_1x_4 and x_4x_2 since by hypothesis E_{12} is normal on every sub-interval and hence E_{13} and E_{32} are both normal. It follows readily that the composite arc $E_{14} + D_{43} + E_{32}$ would have the multipliers of the extremal E_{14} along E_{14}, the multipliers of the extremal tangent to D_{43} at each point of that arc, and the multipliers of the extremal E_{12} along E_{32}. Hence on the composite arcs near E_{12} the value of R would be everywhere different from zero as on E_{12}, and by the differentiability condition of page 684, each such arc would necessarily be an extremal. The extremal E_{12} is, however, the only one having its values y_i, v_i at $x = x_2$, or what is the same thing, its values y_i, y_i', λ_α at $x = x_2$. Hence the arcs $E_{14} + D_{43} + E_{32}$ can not all be minimizing arcs since otherwise all of them and the envelope D would necessarily fall upon E_{12} and their multipliers would coincide with those of E_{12}. But this is impossible because the derivatives $a_k'(t)$, $b_k'(t)$ of the family as determined on page 720 do not all vanish.

If an arc $E_{14} + D_{43} + E_{32}$ is not a minimizing arc it is always possible to find a neighboring admissible arc which joins the points 1 and 2 and gives the integral I a smaller value than $I(E_{14} + D_{43} + E_{32})$, that is, a smaller value than $I(E_{12})$, and hence $I(E_{12})$ can not be a minimum.

The preceding proof of the necessary condition of Mayer is a very satisfactory one geometrically because it emphasizes the geometrical interpretation of the conjugate point and the envelope theorem. But it rests upon two restrictive assumptions, namely, the non-vanishing of the derivative Δ_x at the conjugate point 3, and the requirement that the envelope have a branch projecting from 3 toward 1. In the following sections a proof of an entirely different sort is given which is free from these disadvantages.

25. *The second variation for a normal extremal.* It has been proved on page 17 that if the functions $\eta_i(x)$ of a set of admissible variations for a normal extremal arc E_{12} satisfy the relations $\eta_i(x_1) = \eta_i(x_2) = 0$, then there is a one-parameter family of admissible arcs

$$y_i = y_i(x, b) \qquad (x_1 \leq x \leq x_2)$$

joining the points 1 and 2, containing E_{12} for the parameter value $b = 0$, and having the functions $\eta_i(x)$ as its variations along E_{12}. When the various members of the equations

$$I(b) = \int_{x_1}^{x_2} f[x, y(x,b), y'(x,b)]\,dx,$$
$$0 = \phi_a[x, y(x,b), y'(x,b)]$$

are differentiated for b it is found that

$$I'(b) = \int_{x_1}^{x_2} (f_{y_i} y_{ib} + f_{y_i'} y_{ib}')\,dx,$$

$$0 = \phi_{ay_i} y_{ib} + \phi_{ay_i'} y_{ib}',$$

and a second differentiation gives for $b = 0$

$$I''(0) = \int_{x_1}^{x_2} (f_{y_i} y_{ibb} + f_{y_i'} y_{ibb}' + f_{y_i y_k}\eta_i\eta_k + 2f_{y_i y_k'}\eta_i\eta_k' + f_{y_i' y_k'}\eta_i'\eta_k')\,dx,$$
$$0 = \phi_{ay_i} y_{ibb} + \phi_{ay_i'} y_{ibb}' + \phi_{ay_i y_k}\eta_i\eta_k + 2\phi_{ay_i y_k'}\eta_i\eta_k' + \phi_{ay_i' y_k'}\eta_i'\eta_k'.$$

When the last equations are multiplied by the factors λ_a, integrated from x_1 to x_2, and added to $I''(0)$ this derivative is found to have the value

$$I''(0) = \int_{x_1}^{x_2} (F_{y_i} y_{ibb} + F_{y_i'} y_{ibb}' + 2\omega)\,dx \tag{107}$$

where

$$2\omega(x, \eta, \eta') = F_{y_i y_k}\eta_i\eta_k + 2F_{y_i y_k'}\eta_i\eta_k' + F_{y_i' y_k'}\eta_i'\eta_k'. \tag{108}$$

On account of the equations

$$(d/dx)F_{y_i'} = F_{y_i}$$

the first two terms in the integral (107) have the anti-derivative $F_{y_i'} y_{ibb}$ and this vanishes at x_1 and x_2 as one readily sees by differentiating the equations

$$y_{i1} = y_i(x_1, b), \qquad y_{i2} = y_i(x_2, b)$$

twice with respect to b. Hence the following conclusions are justified:

Along a normal extremal arc E_{12} the second variation of the integral I is always expressible in the form

$$I''(0) = \int_{x_1}^{x_2} 2\omega(x, \eta, \eta')\, dx$$

where 2ω is the quadratic form defined by equation (108). *If $I(E_{12})$ is a minimum for the Lagrange problem then this second variation must be $\geqq 0$ for every set of admissible variations $\eta_i(x)$ whose functions satisfy the relations*

(109) $$\eta_i(x_1) = \eta_i(x_2) = 0.$$

Since admissible variations satisfy the differential equations of variations

(110) $$\Phi_a(x, \eta, \eta') = \phi_{a y_i} \eta_i + \phi_{a y_i'} \eta_i' = 0$$

it is clear that these properties of the second variation suggest a minimum problem in $x\eta$-space of the same type as the original Lagrange problem in xy-space. There is an integral $I''(0)$ which must be $\geqq 0$ in the class of arcs $\eta_i = \eta_i(x)$ in $x\eta$-space satisfying the differential equations (110) and passing through the two fixed points $(x, \eta_1, \cdots, \eta_n) = (x_1, 0, \cdots, 0)$ and $(x, \eta_1, \cdots, \eta_n) = (x_2, 0, \cdots, 0)$, as indicated by equations (109). Evidently the minimum of $I''(0)$ in this class of arcs must be $\geqq 0$ if E_{12} is to be a solution of the original Lagrange problem.

The differential equations of the extremal arcs for the problem in $x\eta$-space are the equations

(111) $$(d/dx)\Omega_{\eta_i'} = \Omega_{\eta_i}, \qquad \Phi_a(x, \eta, \eta') = 0$$

where Ω is a function of the form

(112) $$\Omega(x, \eta, \eta', \mu) = \mu_0 \omega + \mu_a \Phi_a.$$

These are called by von-Escherich [31, Vol. 107, p. 1236] the *accessory system* of linear differential equations. They are the analogues of the Jacobi differential equation for the simplest problem in the plane. If the arc E_{12} is a normal extremal arc for the original Lagrange problem, then every extremal arc for the new problem in $x\eta$-space has this property, since the equations of variation of the linear equations $\Phi_a = 0$ for the $x\eta$-problem are these equations

themselves. Hence it is proper when E_{12} is normal to set $\mu_0 = 1$, the multipliers $\mu_0 = 1$, $\mu_a(x)$ for an extremal arc of the $x\eta$-problem being then unique.

The quadratic form $\Omega(x, \eta, \eta', \mu)$ has the properties

$$2\Omega = \eta_i \Omega_{\eta_i} + \eta_i' \Omega_{\eta_i'} + \mu_a \phi_{\mu_a}, \tag{113}$$

$$u_i \Omega_{v_i} + u_i' \Omega_{v_i'} + \rho_a \Omega_{\sigma_a} = v_i \Omega_{u_i} + v_i' \Omega_{u_i'} + \sigma_a \Omega_{\rho_a}, \tag{114}$$

where the derivatives of Ω are understood to have the arguments (η, η', μ), (u, u', ρ), or (v, v', σ) as indicated by their subscripts. These are well-known formulas for quadratic forms which are readily provable and which will be useful in the following paragraphs.

A final remark concerning the accessory differential equations (111) is also important. These equations are linear and homogeneous in the variables $\eta_i, \eta_i', \eta_i'', \mu_a, \mu_a'$, and the determinant of coefficients of the variables η_i'', μ_a' is the determinant R which will be assumed different from zero along E_{12}. The arguments of Section 6 therefore tell us at once that the accessory equations have one and but one solution η_i, μ_a taking prescribed values of $\eta_i, \Omega_{\eta_i'}$ at a given value of x, or, what is the same thing, prescribed values of η_i, η_i', μ_a satisfying the equations of variation. In particular the only solution taking the values $\eta_i = \Omega_{\eta_i'} = 0$, or $\eta_i = \eta_i' = \mu_a = 0$, at a given x is the set of functions $\eta_i(x) \equiv \mu_a(x) \equiv 0$ which one readily sees to be a solution since the accessory equations are linear and homogeneous in $\eta_i, \eta_i', \eta_i'', \mu_a, \mu_a'$.

26. *A second proof of the analogue of Jacobi's condition.* Consider now a minimizing arc E_{12} for the original Lagrange problem, which has no corners and along which the determinant R of page 684 is everywhere different from zero. According to the differentiability condition on that same page the arc E_{12} must then be an extremal as defined in section 6. For the developments of the present section the additional assumption will be made that the extremal E_{12} is normal on every sub-interval of x_1x_2.

DEFINITION OF CONJUGATE POINT. A value x_3 is said to define a *point 3 conjugate to 1 on the arc* E_{12} if there exists an extremal $\eta_i = u_i(x)$, $\mu_a = \rho_a(x)$ for the $x\eta$-problem whose functions $u_i(x)$ satisfy the relations $u_i(x_1) = u_i(x_3) = 0$ but are not identically zero on x_1x_3. We shall presently see that the definition of a conjugate point on page 720 is equivalent to the one here given.

With this definition agreed upon the necessary condition of Mayer as stated on page 722 can be proved by showing that if there exists a point 3 conjugate to 1 between 1 and 2 on E_{12} then there exists also an admissible

set of variations $\eta_i(x)$ making $I''(0) < 0$. As a first step consider the functions $\eta_i(x)$, $\mu_a(x)$ defined by the equations

$$(115) \qquad \begin{array}{lll} \eta_i(x) \equiv u_i(x), & \mu_a(x) \equiv \rho_a(x) & \text{on } x_1 \leq x \leq x_3, \\ \eta_i(x) \equiv 0, & \mu_a(x) \equiv 0 & \text{on } x_3 \leq x \leq x_2, \end{array}$$

where the functions $u_i(x)$, $\rho_a(x)$ are those indicated in the definition just given for the conjugate point. With the help of the equations (112), (111), (113) it follows readily that for these functions $\eta_i(x)$

$$\begin{aligned} I''(0) &= \int_{x_1}^{x_3} 2\omega(x, \eta, \eta')\,dx = \int_{x_1}^{x_3} 2\Omega(x, u, u', \rho)\,dx \\ &= \int_{x_1}^{x_3} (u_i\Omega_{u_i} + u_i'\Omega_{u_i'} + \rho_a\Omega_{\rho_a})\,dx \\ &= u_i\Omega_{u_i'} \Big|_1^3 = 0. \end{aligned}$$

The functions $\eta_i(x)$ in (115) can not minimize $I''(0)$, however, since, as will be shown in the next paragraph, they do not satisfy the corner conditions

$$(116) \qquad \Omega_{\eta_i'}\ [x, \eta, \eta'(x-0), \mu(x-0)] = \Omega_{\eta_i'}\ [x, \eta, \eta'(x+0), \mu(x+0)]$$

at the point x_3. Hence there must be other admissible variations $\eta_i(x)$ vanishing at x_1 and x_2 and giving $I''(0)$ a value less than zero, and $I(E_{12})$ can not be a minimum.

To show that the corner conditions are not satisfied one may calculate readily the values of the derivatives $\Omega_{\eta_i'}$ for the functions (115) at the left and right of x_3. It is found then that the corner conditions (116) would require that $\Omega_{u_i'} = 0$ at the point x_3 as well as $u_i = 0$, and according to a remark at the end of the preceding section the functions $u_k(x)$, $\rho_a(x)$ would then have to be identically zero, which is not the case. The proof of Mayer's condition is now complete.

27. *The determination of conjugate points.* For a one-parameter family of extremals

$$y_i = y_i(x, b), \qquad \lambda_a = \lambda_a(x, b)$$

the equations

$$(d/dx)F_{y_i'} = F_{y_i}, \qquad \phi_a = 0$$

are identities in x and b. When they are differentiated with respect to b we find

$$\begin{gathered} (d/dx)(F_{y_i'y_k}y_{kb} + F_{y_i'y_k'}y_{kb}' + F_{y_i'\lambda_a}\lambda_{ab}) = F_{y_iy_k}y_{kb} + F_{y_iy_k'}y_{kb}' + F_{y_i\lambda_a}\lambda_{ab}, \\ \phi_{ay_k}y_{kb} + \phi_{ay_k'}y_{kb}' = 0, \end{gathered}$$

and these are precisely the accessory equations with the arguments $\eta_i = y_{ib}$, $\mu_a = \lambda_{ab}$. A $2n$-parameter family of extremals defined by equations similar to equations (35) or (37) on pages 686-7 furnishes by this differentiation process $2n$ solutions

$$(117) \qquad \begin{array}{l} y_{1a_k}, \cdots, y_{na_k};\ \lambda_{1a_k}, \cdots, \lambda_{ma_k} \\ y_{1b_k}, \cdots, y_{nb_k};\ \lambda_{1b_k}, \cdots, \lambda_{mb_k} \end{array} \qquad (k = 1, \cdots, n)$$

of the accessory equations. The formulas of most importance here are those for $2n$-parameter families for which the determinant (36) is different from zero at some point, say x_1. We shall see in the next paragraph that it is then different from zero for all values of x.

Since the determinant R is different from zero along E_{12} the equations

$$\zeta_i = \Omega_{\eta_i'}(x, \eta, \eta', \mu), \qquad \Phi_a(x, \eta, \eta') = 0,$$

analogous to equations (30) on page 685, can be solved for η_k', μ_ρ. The solution has the form

$$(118) \qquad \eta_k' = G_k(x, \eta, \zeta), \qquad \mu_\beta = H_\beta(x, \eta, \zeta),$$

and the accessory equations are equivalent to the equations

$$(119) \qquad (d\eta_k/dx) = G_k(x, \eta, \zeta),\ (d\zeta_k/dx) = \Omega_{\eta_k}(x, \eta, G(x, \eta, \zeta), H(x, \eta, \zeta)).$$

All of these equations are linear and homogeneous in the arguments η_i, η_i', μ_a, ζ_i where they occur. For equations of the type (119) it is well known * that $2n$ solutions (η_k, ζ_k) whose determinant is different from zero at a single value of x, will have that determinant different from zero for all values x, and that every other solution is linearly expressible with constant coefficients in terms of $2n$ solutions which have this property. Every solution of the accessory equations is therefore expressible linearly with constant coefficients in terms of the $2n$ corresponding sets (η_k, μ_β) defined by the second of equations (118).*

Since the determinant (36) of page 687 is different from zero at $x = x_1$ it follows that it is different from zero for all values of x. For the $2n$ solutions (117) of the accessory equations define $2n$ solutions (η_k, ζ_k) of equations (119) whose determinant is different from zero. Hence every solution $(\eta_i, \mu_a) = (u_i, \rho_a)$ of the accessory equations is expressible in the form

$$u_i = c_k y_{ia_k} + d_k y_{ib_k}, \quad \rho_a = c_k \lambda_{aa_k} + d_k \lambda_{ab_k}.$$

---

* See, for example, Goursat, *A Course in Mathematical Analysis*, translated by Hedrick and Dunkel, Vol. 2, Part 2, pp. 153-4.

The values x_3 determining conjugate points according to the definition on page 725 are those for which the equations

$$u_i(x_3) = c_k y_{ia_k}(x_3) + d_k y_{ib_k}(x_3) = 0,$$
$$u_i(x_1) = c_k y_{ia_k}(x_1) + d_k y_{ib_k}(x_1) = 0,$$

have solutions c_k, d_k not all zero. But these are precisely the values x_3 for which the determinant $\Delta(x, x_1, a, b)$ vanishes, as indicated in the definition on page 720. We shall see on page 740 that for every ξ on x_1x_2 the zeros of $\Delta(x, \xi, a, b)$ are isolated from ξ when an extension of E_{12} is normal on every sub-interval.

Consider now an n-parameter family of extremals

$$y_i = y_i(x, b_1, \cdots, b_n), \quad \lambda_a = \lambda_a(x, b_1, \cdots, b_n)$$

all of which pass through the point 1, and such that the functions $v_i = F_{y_i'}$ for the family have their determinant $| v_{ib_k} |$ different from zero at x_1. All of the derivatives y_{ib_k} vanish at x_1 as one may see by differentiating the equations

$$y_{i1} = y_i(x_1, b_1, \cdots, b_n)$$

with respect to b_k. Every solution η_i, μ_a of the accessory equations for which the functions $\eta_i(x)$ all vanish at x_1 is expressible in the form

$$\eta_i = c_k y_{ib_k}, \quad \mu_a = c_k \lambda_{ab_k},$$

where the coefficients c_k are constants. For such a solution is uniquely determined by its set of values $\eta_i = 0$, $\zeta_i = \Omega_{\eta_i'}$ at $x = x_1$. If the constants c_k are solutions of the equations

$$\zeta_i(x_1) = c_k v_{ib_k}(x_1),$$

which in fact determine them uniquely, then the two solutions η_i, μ_a and $c_k y_{ib_k}$, $c_k \lambda_{ab_k}$ of the accessory equations have the same values $\eta_i = 0$, ζ_i at $x = x_1$ and hence are identical for all values of x. It follows that the points 3 conjugate to 1 on E_{12} are determined by values x_3 for which the equations

$$c_k y_{ib_k}(x_3) = 0$$

have solutions c_k not all zero, that is, by values $x_3 \neq x_1$ which make the determinant $D(x, b) = | y_{ib_k} |$ vanish. These results may be summarized as follows:

Let E_{12} be an extremal arc which is contained in a $2n$-parameter family of extremals

$$y_i = y_i(x, a_1, \cdots, a_n, b_1, \cdots, b_n), \quad \lambda_a = \lambda_a(x, a_1, \cdots, a_n, b_1, \cdots, b_n)$$

for special values a_{i0}, b_{i0} *of the parameters. Suppose furthermore that the determinant*

$$\begin{vmatrix} y_{ia_k} & y_{ib_k} \\ v_{ia_k} & v_{ib_k} \end{vmatrix}$$

of the family, where $v_i = F_{y_i'}(x, y, y', \lambda)$, *is different from zero at the point 1 on* E_{12}. *Then the points 3 conjugate to 1 on* E_{12} *are determined by the roots* $x_3 \neq x_1$ *of the function* $\Delta(x, x_1, a_0, b_0)$ *where*

$$\Delta(x, x_1, a, b) = \begin{vmatrix} y_{ia_k}(x, a, b) & y_{ib_k}(x, a, b) \\ y_{ia_k}(x_1, a, b) & y_{ib_k}(x_1, a, b) \end{vmatrix}.$$

If E_{12} *is a member of an n-parameter family of extremals*

$$y_i = y_i(x, b_1, \cdots, b_n), \quad \lambda_a = \lambda_a(x, b_1, \cdots, b_n)$$

all of which pass through the point 1, and such that the determinant $|v_{ib_k}|$ *for the functions* $v_i = F_{y_i'}$ *belonging to the family is different from zero at the point 1 on* E_{12}, *then the points conjugate to 1 on* E_{12} *are determined by the roots* $x_3 \neq x_1$ *of the function* $D(x, b_0)$ *where*

$$D(x, b) = |y_{ib_k}|$$

and the b_{i0} *are the parameter values defining* E_{12}.

CHAPTER IV.

SUFFICIENT CONDITIONS FOR A MINIMUM.

The conditions developed in the preceding chapters are conditions which must be satisfied by every minimizing arc for the Lagrange problem, but they have not been shown to actually insure the minimizing property. In this chapter it is proposed to discuss sets of conditions which are sufficient for a minimum. The methods of proof used are in essence those which Weierstrass applied in similar cases and which have been extended to the Lagrange problem by A. Mayer, Bolza, and others, but they involve important simplifications and improvements.

28. *Mayer fields and the fundamental sufficiency theorem.* The notion of a field has been defined in a number of different ways. The definition given here is not the usual one and is somewhat sophisticated, but it emphasizes properties which are well known for fields of the simplest problem in the

plane, and leads promptly to the theorem which is fundamental for all of the sufficiency proofs. In order to phrase this definition as simply as possible let us agree to call a set of values (x, y, y') *admissible* if it lies interior to the region $\mathfrak{R}$ where the continuity properties of the functions f and ϕ_a have been assumed, and satisfies the equations $\phi_a = 0$, and gives the matrix $\| \phi_{a y_i'} \|$ the rank m.

DEFINITION OF A MAYER FIELD. A *Mayer field* is a region $\mathfrak{F}$ of xy-space containing only interior points and having associated with it a set of functions

$$p_i(x, y), \quad l_a(x, y)$$

with the following properties:

(a) they have continuous first partial derivatives in $\mathfrak{F}$;
(b) the sets $(x, y, p(x, y))$ defined by the points (x, y) in $\mathfrak{F}$ are all admissible;
(c) the integral

$$I^* = \int \{F(x, y, p, l)\,dx + (dy_i - p_i dx) F_{y_i'}(x, y, p, l)\}$$

formed with these functions is independent of the path in $\mathfrak{F}$.

The integral I^* can also be written in the form

$$I^* = \int \{A dx + B_i dy_i\}$$

where

$$A(x, y) = F(x, y, p, l) - p_k F_{y_k'}(x, y, p, l),$$
$$B_i(x, y) = F_{y_i'}(x, y, p, l).$$

If such an integral is independent of the path every arc is a minimizing arc for it and the Euler-Lagrange differential equations applied to it give the well-known conditions

$$\partial A/\partial y_i = \partial B_i/\partial x, \quad \partial B_i/\partial y_k = \partial B_k/\partial y_i \tag{120}$$

as necessary conditions for its invariantive property. One may readily prove the identities

$$\begin{aligned} \partial A/\partial y_i - \partial B_i/\partial x = F_{y_i} - (\partial/\partial x) F_{y_i'} - p_k(\partial/\partial y_k) F_{y_i'} \\ + p_k(\partial B_i/\partial y_k - \partial B_k/\partial y_i) + \phi_a\, \partial l_a/\partial y_i \end{aligned} \tag{121}$$

where the partial derivatives indicated by the symbols ∂ are taken with respect to the independent variables x, y_i which occur explicitly and also in the field functions $p_i(x, y)$, $l_a(x, y)$.

From these results it is easy to see that in the field $\mathfrak{F}$ every solution $y_i(x)$ of the equations

$$(122) \qquad dy_i/dx = p_i(x, y)$$

is an extremal with the multipliers $\lambda_a = l_a(x, y(x))$. For in the first place such an arc necessarily satisfies the equations $\phi_a = 0$, since the values (x, y, p) are all admissible; and in the second place the equations (120) and (121) then show that along such an arc

$$F_{y_i} - (d/dx)F_{y_i'} = F_{y_i} - (\partial/\partial x)F_{y_i'} - p_k(\partial/\partial y_k)F_{y_k'} = 0.$$

The arcs satisfying equations (122) are called the *extremals of the field.* Through each point of $\mathfrak{F}$ there passes one and but one such extremal arc since the equations (122) are of the first order. Furthermore the value of I^* along an extremal arc of the field is equal to that of the original integral I, since the equations $dy_i - p_i dx = 0$ are all satisfied along the field extremals.

If E_{12} is an extremal arc of a field $\mathfrak{F}$ then for every admissible arc C_{12} in the field joining the same two points 1 and 2 the formula

$$(123) \qquad I(C_{12}) - I(E_{12}) = \int_{x_1}^{x_2} E[x, y, p(x, y), y', l(x, y)]dx$$

holds, where

$$E = F(x, y, y', l) - F(x, y, p, l) - (y_i' - p_i)F_{y_i'}(x, y, p, l)$$

and the arguments $y(x)$, $y'(x)$ in the integrand are those belonging to C_{12}.

The formula (123) is the analogue of a well-known one of Weierstrass and the proof of it is very simple. For since I^* is independent of the path in $\mathfrak{F}$ and has the same values as I along an extremal of the field it follows that

$$I(E_{12}) = I^*(E_{12}) = I^*(C_{12}),$$

and hence that

$$I(C_{12}) - I(E_{12}) = I(C_{12}) - I^*C_{12}).$$

The last two terms give the integral in the second member of the formula (123) when the integrand f in $I(C_{12})$ is replaced by F. This is evidently permissible since C_{12} is by hypothesis an admissible arc and therefore satisfies the equations $\phi_a = 0$.

With these results in mind it is now possible to prove the following important theorem:

THE FUNDAMENTAL SUFFICIENCY THEOREM. *If E_{12} is an extremal arc of a field $\mathfrak{F}$ and if at each point of the field the condition*

$$E[x, y, p(x, y), y', l(x, y)] > 0$$

holds for every admissible set (x, y, y') *different from* (x, y, p), *then the inequality* $I(C_{12}) > I(E_{12})$ *is true for every admissible arc* C_{12} *in the field and joining the end-points of* E_{12} *but not identical with* E_{12}.

It is evident from formula (123) that the inequality $I(C_{12}) \geqq I(E_{12})$ is necessarily satisfied. The equality sign is appropriate only if the E-function vanishes at every point of C_{12}, that is, only if the equations $y_i' = p_i$ are satisfied at each point of C_{12}. But in that case the arc C_{12} would coincide with E_{12} since the equations $y_i' = p_i$ have only one solution through the point 1 and that is E_{12} itself.

29. *The construction of a field.* The extremal arcs of a field may be regarded as forming an n-parameter family since one of them passes through each point of the field. By analogy with the properties of fields for the simplest problem of the calculus of variations in the plane it might be expected that every n-parameter family of extremals which simply covers a region in xy-space would provide a set of slope functions and multipliers $p_i(x, y)$, $l_\alpha(x, y)$ which would make the integral I^* independent of the path in that region, and hence form a field over the region, but such is not the case. The n-parameter families which can form fields are special in character in somewhat the same way that a two-parameter family of straight lines in xyz-space is special if it is cut orthogonally by a surface. It is well known that not every such family of straight lines has an orthogonal surface.

Let the equations

$$(124) \qquad y_i = y_i(x, a_1, \cdots, a_n), \quad \lambda_\alpha = \lambda_\alpha(x, a_1, \cdots, a_n)$$

be an n-parameter family of extremals with the property that the functions y_i, y_{ix}, λ_α have continuous first partial derivatives for all values $(x, a_1, \cdots, a_n)$ satisfying conditions of the form

$$(125) \qquad \begin{gathered} \xi_1(a_1, \cdots, a_n) \leqq x \leqq \xi_2(a_1, \cdots, a_n). \\ (a_1, \cdots, a_n) \text{ in a region } A. \end{gathered}$$

Suppose further that there is an n-space

$$x = x_1(a_1, \cdots, a_n), \quad y_i = y_i(x_1(a_1, \cdots, a_n), a_1, \cdots, a_n)$$

cutting the extremals (124) for which the function $x_1(a_1, \cdots, a_n)$ has continuous first partial derivatives in A. The extremals (124) are said to simply cover a field $\mathfrak{F}$ of points (x, y) if to each point of the region there corresponds one and but one set of values x, $a_i(x, y)$ satisfying the first n equations (124)

and the conditions (125), and if the functions $a_i(x,y)$ so defined have continuous derivatives in $\mathfrak{F}$. The functions

$$p_i(x,y) = y_{ix}[x, a(x,y)], \quad l_a(x,y) = \lambda_a[x, a(x,y)]$$

are then a set of slope-functions and multipliers for the region $\mathfrak{F}$, and the following theorem can be proved:

Suppose that an n-parameter family of extremals

$$y_i = y_i(x, a_1, \cdots, a_n), \quad \lambda_a = \lambda_a(x, a_1, \cdots, a_n) \tag{126}$$

is intersected by an n-space

$$x = x_1(a_1, \cdots, a_n), \quad y_i = y_i(x_1(a_1, \cdots, a_n), a_1, \cdots, a_n) \tag{127}$$

and simply covers a region $\mathfrak{F}$ of xy-space containing only interior points, in the manner described in the preceding paragraphs. If the parameter values of the extremal through a point (x,y) are denoted by $a_i(x,y)$ then the region $\mathfrak{F}$ is a field with the slope-functions and multipliers

$$p_i(x,y) = y_{ix}[x, a(x,y)], \quad l_a(x,y) = \lambda_a[x, a(x,y)] \tag{128}$$

provided that the integral I^* is independent of the path in the n-space (127).

The proof may be made with the help of the Auxiliary Theorem II of page 716. For an arc D_{46} in $\mathfrak{F}$ with equations of the form

$$x = x(t), \quad y_i = y_i(t) \qquad (t' \leq t \leq t'')$$

defines a one-parameter family of extremals intersecting it, and a corresponding arc C_{35} in the n-space (127), by means of the functions $a_i(t) = a_i[x(t), y(t)]$. According to the auxiliary theorem cited it is then true that

$$I^*(D_{46}) = I^*(C_{35}) + I(E_{56}) - I(E_{34}).$$

The three terms on the right are completely determined when the end-points of D_{46} are given, since by hypothesis the value $I^*(C_{35})$ is the same for all arcs C_{35} with the same end-points in the n-space (127). Hence the integral I^* is independent of the path in the whole of the region $\mathfrak{F}$, as required by the definition of a field.

The preceding theorem suggests at once a number of methods of constructing fields by means of n-parameter families of extremals. One may take the n-parameter family through a fixed point O and regard the point O as a degenerate n-space (127). Certainly on this degenerate n-space the integral I^* is independent of the path. Every region in xy-space simply covered by

the extremals will then be a field with the slope-functions and multipliers (128).

If an n-space

$$(129) \qquad x = X(a_1, \cdots, a_n), \quad y_i = Y_i(a_1, \cdots, a_n)$$

and a function $W(a_1, \cdots, a_n)$ are chosen arbitrarily in advance the $n + m$ equations

$$(130) \qquad FX_{a_i} + (Y_{ka_i} - y_k' X_{a_i}) F_{y_k'} = W_{a_i}, \quad \phi_a = 0,$$

where the arguments of F, ϕ_a are X, Y_i, y_i', λ_a, may under certain conditions be solved for the $n + m$ variables y_i', λ_a as functions of $a_1, \cdots, a_n$. At each point of the n-space an initial element x, y_i, y_i', λ_a of an extremal is thus determined, and the extremals which have these initial elements form an n-parameter family. The integrand of the integral I^* for this family has the value dW on every arc in the n-space (129), on account of the equations (130), since along such an arc the differentials dx, dy_k have the values

$$dx = X_{a_i} da_i, \quad dy_k = Y_{ka_i} da_i.$$

Hence the integral I^* will be independent of the path on the space (129) and every region of xy-space simply covered by the family of extremals will form a field. If the derivatives W_{a_i} all vanish then an n-space (129) which satisfies the equations (130) with the extremals of the family it is said to cut the family transversally.

A similar discussion can be made for initial spaces (129) of lower dimensions.

30. *Sufficient conditions for a strong relative minimum.* In the following paragraphs the necessary conditions deduced in the preceding chapters will be designated by the numerals I, II, III, IV. These are, respectively, the necessary condition of page 683, the analogue of Weierstrass' condition on page 718, the condition of Clebsch on page 719, and the condition of Mayer on page 722. The notations II′, III′ will be used to designate the conditions II and III when strengthened to exclude the equality sign which occurs in their statements. Similarly IV′ is the stronger condition of Mayer which excludes the conjuate point 3 from the end-point 2 of E_{12}, as well as from the interior of that arc. An arc E_{12} with multipliers $\lambda_0 = 1$, $\lambda_a(x)$ will be said to satisfy the condition II_b' if the inequality

$$E(x, y, y', Y', \lambda) > 0$$

holds for every set of elements (x, y, y', Y', λ) for which the set (x, y, y', λ)

is in a neighborhood of similar sets belonging to E_{12}, and $(x, y, Y') \neq (x, y, y')$ is admissible.

Every extremal arc E_{12} defined on an interval $x_1 x_2$, and on which the determinant R is different from zero, defines an extended extremal on an interval $x_1 - d \leqq x \leqq x_2 + d$ which contains E_{12} as part of it. We may call this longer extremal an extension of E_{12}.

With these agreements we can state the following theorem:

SUFFICIENT CONDITIONS FOR A STRONG RELATIVE MINIMUM. *If an admissible arc E_{12}, without corners and with an extension normal on every subinterval, satisfies the conditions I, II_b', III', IV', then there is a neighborhood $\mathfrak{F}$ of the points (x, y) on E_{12} such that the inequality $I(C_{12}) > I(E_{12})$ holds for every admissible arc C_{12} which is in $\mathfrak{F}$ and not identical with E_{12}.*

The minimum furnished by E_{12} is called a relative minimum because it is in a class of arcs restricted to lie in a neighborhood $\mathfrak{F}$ of E_{12}; and it is a strong relative minimum because the neighborhood $\mathfrak{F}$ lays no restriction on the slopes y_i' of comparison arcs which lie in it.

In order to prove the theorem we should note in the first place that the condition *I* and the normality of E_{12} imply a unique set of multipliers $\lambda_0 = 1$, $\lambda_a(x)$ and constants c_i with which E_{12} satisfies the equations (24) of page 683.

The condition *III'* now implies that the determinant R of page 684 is different from zero at every element (x, y, y', λ) of E_{12}. For at an element where R vanished the linear equations

$$F_{y_i' y_k'} \Pi_k + \phi_{a y_i'} \mu_a = 0, \quad \phi_{a y_k'} \Pi_k = 0 \tag{131}$$

would have solutions Π_k, μ_a not all zero, with the numbers Π_k also not all zero since the matrix $\| \phi_{a y_i'} \|$ has rank m. But when the first equations (131) are multiplied by $\Pi_1, \cdots, \Pi_n$ and added it is found that

$$F_{y_i' y_k'} \Pi_i \Pi_k = 0,$$

as a result of the second set of equations (131), which would contradict the condition III'.

Since the determinant R is different from zero along E_{12} it follows from the differentiability condition of page 684 that E_{12} must be an extremal. According to the developments of Section 6, page 687, there exists a $2n$-parameter family of extremals

$$y_i = y_i(x, a, b), \quad \lambda_a = \lambda_a(x, a, b)$$

containing E_{12} for special parameter values a_{i0}, b_{i0}. The functions y_i, y_{ix}, λ_a have continuous partial derivatives of the first three orders near the values

(x, a_i, b_i) belonging to E_{12}, and the determinant (36) of page 687 is different from zero at the point 1 on E_{12}.

It will be shown in Section 32 that for an arc E_{12} with an extension normal on every sub-interval there is always an interval $x_1 - h \leq x \leq x_1 + h$ containing no pair of conjugate points, or in other words, containing no two values x, x_0 which satisfy the equation $\Delta(x, x_0, a_0, b_0) = 0$, where Δ is the determinant (104) of page 719. Hence if $x_0 < x_1$ be chosen sufficiently near to x_1 the function $\Delta(x, x_0, a_0, b_0)$ will be different from zero on the interval $x_1 \leq x \leq x_1 + h$, and different from zero also in the interval $x_1 + h \leq x \leq x_2$ on account of the continuity of Δ and condition IV′. The equations

$$y_i = y_i(x, a, b), \qquad y_{i0} = y_i(x_0, a, b) \tag{132}$$

have now as initial solutions the totality of values (x, y, a, b) belonging to E_{12}, and their functional determinant $\Delta(x, x_0, a, b)$ with respect to the parameters a_i, b_i is different from zero at these initial solutions on account of the choice of x_0 which has just been made. Well-known implicit function theorems then justify the statement that there is a neighborhood $\mathfrak{F}$ of the points (x, y) on E_{12} in which the equations (132) have solutions $a_i(x, y)$, $b_i(x, y)$ with continuous partial derivatives of the first three orders since the functions (132) have such derivatives. This neighborhood $\mathfrak{F}$ is a field with the slope functions and mutlipliers

$$p_i(x, y) = y_{ix}[x, a(x, y), b(x, y)], \quad \lambda_a(x, y) = \lambda_a[x, a(x, y), b(x, y)]$$

since the extremals which simply cover it all pass through the fixed point 0 corresponding, on E_{12} extended, to the value x_0. If the field $\mathfrak{F}$ is taken sufficiently small the values $x, y, p_i(x, y), \lambda_a(x, y)$ belonging to it will remain in so small a neighborhood of the sets (x, y, y', λ) belonging to E_{12} that according to the condition II_b' the inequality

$$E[x, y, p(x, y), y, \lambda(x, y)] > 0 \tag{133}$$

will hold for every admissible element $(x, y, y') \neq (x, y, p)$ in $\mathfrak{F}$. The fundamental sufficiency theorem then justifies the theorem which was to be proved.

31. *Sufficient conditions for a weak relative minimum.* The conditions I, III′, IV′ were the only ones used in the last section up to the very last paragraph. If they only are assumed it is not possible to establish the condition (133). The E-function for admissible elements (x, y, y') in the field $\mathfrak{F}$ is expressible, however, with the help of Taylor's formula with integral remainder term, in the form

$$E = (y_i' - p_i)(y_k' - p_k) \int_0^1 (1 - \theta) F_{y_i' y_k'}[x, y, p + \theta(y' - p), \lambda] d\theta \tag{134}$$

where $p_i = p_i(x, y)$, $\lambda_a = \lambda_a(x, y)$ are the slope-functions and multipliers of the field, and the differences $y_i' - p_i$ satisfy the equation

$$\phi_a(x, y, y') - \phi_a(x, y, p) = (y_i' - p_i) \int_0^1 \phi_{a y_i'}[x, y, p + \theta(y' - p)] d\theta = 0.$$

On account of the condition III′ the quadratic form

$$\Pi_i \Pi_k \int_0^1 (1 - \theta) F_{y_i' y_k'}[x, y, p + \theta(y' - p), \lambda] d\theta$$

is positive for all sets (x, y, y', Π) for which (x, y, y') is on the arc E_{12} where $y_i' = p_i(x, y)$, and for which the numbers Π_i satisfy the equations

$$\Pi_i \Pi_i = 1, \quad \Pi_i \int_0^1 \phi_{a y_i'}[x, y, p + \theta(y' - p)] d\theta = 0.$$

Hence it stays positive for sets of values (x, y, y', Π) for which the numbers Π_i satisfy these equations and the set (x, y, y') lies in a sufficiently small neighborhood N of similar sets on E_{12}. It follows readily that the E-function (134) of the field $\mathfrak{F}$ is positive at least for all sets $(x, y, y') \neq (x, y, p)$ in the neighborhood N, and the following theorem is therefore justified:

SUFFICIENT CONDITIONS FOR A WEAK RELATIVE MINIMUM. *If an admissible arc E_{12} without corners and with an extension normal on every sub-interval, satisfies the conditions I, III′, IV′ then there is a neighborhood N of the sets of values (x, y, y') on E_{12} such that the inequality $I(C_{12}) > I(E_{12})$ holds for every admissible arc C_{12} whose elements (x, y, y') are all in N but which is not identical with E_{12}.*

The minimum described in this theorem is called a weak relative minimum because the neighborhood N in which it exists requires the slopes y_i' of the comparison arcs C_{12}, as well as their points (x, y), to be near those on E_{12}.

32. *The justification of a preceding statement.* It was stated on page 736 that there is always an interval $x_1 - h \leq x \leq x_1 + h$ on which no two values x, x_0 can satisfy the equation $\Delta(x, x_0, a_0, b_0) = 0$. The proof of this statement is not simple, but it can be made with the help of properties of solutions of the accessory differential equations

$$(135) \qquad (d/dx)\Omega_{\eta_i'} - \Omega_{\eta_i} = 0, \quad \Omega_{\mu_a} = \Phi_a = 0$$

for the arc E_{12} described on page 724. It is understood that the arc E_{12} is an extremal with an extension normal on every sub-interval and satisfying the condition III′. As a consequence of these properties the determinant R is different from zero at every point of E_{12}.

The equation (114) on page 725

$$u_i\Omega_{v_i} + u_i'\Omega_{v_i'} + \rho_a\Omega_{\sigma_a} = v_i\Omega_{u_i} + v_i'\Omega_{u_i'} + \sigma_a\Omega_{\rho_a},$$

justifies readily the further relation

$$u_i[\Omega_{v_i} - (d/dx)\Omega_{v_i'}] + \rho_a\Omega_{\sigma_a} - v_i[\Omega_{u_i} - (d/dx)\,\Omega_{u_i'}] - \sigma_\sigma\Omega_{\rho_a}$$
$$= (d/dx)(v_i\Omega_{u_i'} - u_i\Omega_{v_i'}).$$

Hence for every pair of solutions u_i, ρ_a and v_i, σ_a of the accessory equations the expression

$$\psi(u, \rho, v, \sigma) = u_i\Omega_{v_i'} - v_i\Omega_{u_i'}$$

is a constant. If this constant is zero the two solutions are said to be *conjugate solutions.*

There is one and but one set of solutions η_i, μ_a of the accessory equations (135) for which η_i, $\zeta_i = \Omega_{\eta_i'}$ take assigned values at the value x_1, as shown for the original xy-problem on pages 685 and 686. A matrix of n solutions u_{ik}, ρ_{ak} $(k = 1, \cdots, n)$ therefore exists for which at the value x_1 the matrix $\| u_{ik} \|$ is the identity matrix and the corresponding matrix of the functions $\zeta_i = \Omega_{\eta_i'}$ has all its elements zero. The solutions u_{ik}, ρ_{ak} $(k = 1, \cdots, n)$ are conjugate in pairs, as one readily verifies, since their functions ζ_i all vanish at x_1. The notations u_i, ρ_a and v_i, σ_a will be used for the linear expressions

$$u_i = a_k u_{ik}, \qquad \rho_a = a_k\rho_{ak},$$
$$v_i = a_k' u_{ik}, \qquad \sigma_a = a_k'\rho_{ak},$$

where the coefficients a_k are functions of x to be determined and the variables a_k' are derivatives of the coefficients a_k with respect to x. Primes attached to expressions involving u_i, ρ_a or v_i, σ_a will always indicate derivatives of those expressions with respect to x calculated as if the coefficients a_k, a_k' were independent of x. One readily verifies, then, the relations

$$(136) \qquad \begin{gathered} (\Omega_{u_i'})' = \Omega_{u_i}, \quad (\Omega_{v_i'})' = \Omega_{v_i}, \quad u_i\Omega_{v_i'} - v_i\Omega_{u_i'} = 0, \\ (d/dx)\Omega_{u_i'} = (\Omega_{u_i'})' + \Omega_{v_i'} = \Omega_{u_i} + \Omega_{v_i'} \end{gathered}$$

in which it is understood that the differentiation indicated by d/dx takes account of the fact that the coefficients a_i are functions of x.

Let the functions $\eta_i(x)$ be a set of admissible variations along the arc E_{12}, satisfying therefore the equations $\Phi_a = 0$. The equations

$$\eta_i = u_i = a_k u_{ik}, \quad \mu_a = \rho_a = a_k\rho_{ak}$$

determine uniquely the coefficients a_k and the multipliers μ_a as functions of x on an interval $x_1 - h \leq x \leq x_1 + h$ chosen so small that on it the determinant $|u_{ik}|$ is everywhere different from zero. The derivatives η_i' have the values

$$\eta_i' = a_k u_{ik}' + a_k' u_{ik} = u_i' + v_i. \tag{137}$$

With the help of Taylor's formula, equation (113) of page 725, the equations (136) and (137) above, and the relations $\Omega_{\rho_a} = \Phi_a = 0$, one verifies the further relations

$$\begin{aligned} 2\omega(x, \eta, \eta') &= 2\Omega(x, \eta, \eta', \rho) = 2\Omega(x, u, u' + v, \rho) \\ &= 2\Omega(x, u, u', \rho) + 2v_i\Omega_{u_i'} + F_{y_i' y_k'} v_i v_k \\ &= u_i\Omega_{u_i} + u_i'\Omega_{u_i'} + \rho_a\Omega_{\rho_a} + 2v_i\Omega_{u_i'} + F_{y_i' y_k'} v_i v_k \\ &= u_i[\Omega_{u_i} + \Omega_{v_i'}] + (u_i' + v_i)\Omega_{u_i'} + F_{y_i' y_k'} v_i v_k \\ &= (d/dx)(\eta_i\Omega_{u_i'}) + F_{y_i' y_k'}(\eta_i' - u_i')(\eta_k' - u_k'). \end{aligned}$$

For arbitrary multipliers $\mu_a(x)$ taken with the functions $\eta_i(x)$ it follows therefore that

$$2\Omega(x, \eta, \eta', \mu) = (d/dx)\eta_i\Omega_{u_i'} + F_{y_i' y_k'}(\eta_i' - u_i')(\eta_k' - u_k')$$

and hence with the help of equation (113) on page 725 that

$$\eta_i[\Omega_{\eta_i} - (d/dx)\Omega_{\eta_i'}] + (d/dx)\eta_i(\Omega_{\eta_i'} - \Omega_{u_i'}) = F_{y_i' y_k'}(\eta_i' - u_i')(\eta_k' - u_k').$$

The last equation justifies the following lemma:

LEMMA. *There is an interval $x_1 - h \leq x \leq x_1 + h$ on which there exists no solution $\eta_i(x)$, $\mu_a(x)$ of the accessory equations, except the solution $\eta_i \equiv \mu_a \equiv 0$, whose elements $\eta_i(x)$ all vanish at two points x' and x'' of the interval; or, in other words, there is an interval on which no pair of values x', x'' can define conjugate points on E_{12}.*

This is clear since the last equation shows that for a system of solutions $\eta_i(x)$, $\mu_a(x)$ of the accessory equations the sum $\eta_i(\Omega_{\eta_i'} - \Omega_{u_i'})$ has a non-negative derivative on $x_1 - h \leq x \leq x_1 + h$, on account of the property III′ of E_{12}. If the functions $\eta_i(x)$ all vanish at two points x' and x'' the differences $\eta_i' - u_i' = v_i$ are identically zero on $x'x''$, and this implies that the derivatives a_k' are all zero and the coefficients a_k constants. But since the $\eta_i(x)$ vanish at x' and $|u_{ik}|$ is different from zero these coefficients are then all zero, and the functions $\eta_i(x)$ vanish identically on $x'x''$. The multipliers $\mu_a(x)$ are also zero on $x'x''$. Otherwise they would form with $\lambda_0 = 0$ a set of multipliers for E_{12}, as one readily sees by examining the accessory equations, and this is impossible since the extension of E_{12} is normal on $x'x''$ if the interval $x_1 - h \leq x \leq x_1 + h$ is taken sufficiently small.

As an immediate consequence of this lemma we have the following corollary:

COROLLARY. *There is an interval $x_1 - h \leq x \leq x_1 + h$ on which the determinant*

$$\Delta(x, x_0, a, b) = \begin{vmatrix} y_{ia_k}(x) & y_{ib_k}(x) \\ y_{ia_k}(x_0) & y_{ib_k}(x_0) \end{vmatrix},$$

formed for a family of extremals $y_i = y_i(x, a, b)$, $\lambda_a(x, a, b)$ as described in the theorem of page 687, can not vanish for any pair of points $(x_0, x) = (x', x'')$.

The solutions η_i, μ_a of the accessory equations are all expressible in the form

$$\eta_i = c_k y_{ia_k} + d_k y_{ib_k}, \qquad \mu_a = c_k \lambda_{aa_k} + d_k \lambda_{ab_k}, \tag{138}$$

as was indicated on page 727. If $\Delta(x'', x', a, b) = 0$ for points x', x'' on the interval $x_1 - h \leqq x \leqq x_1 + h$ then there would be constants c_k, d_k not all zero such that the solution (138) has $\eta_i(x') = \eta_i(x'') = 0$, and by the lemma it would follow that $\eta_i \equiv \mu_a \equiv 0$. In that case the corresponding functions

$$\zeta_i = c_k v_{ia_k} + d_k v_{ib_k} = \Omega_{\eta_i'}$$

would also vanish identically, which is impossible since the determinant

$$\begin{vmatrix} y_{ia_k} & y_{ib_k} \\ v_{ia_k} & v_{ib_k} \end{vmatrix}$$

of page 687 is by hypothesis different from zero.

CHAPTER V.

HISTORICAL REMARKS.

A complete history of the problem of Lagrange would require an extensive presentation. The remarks in the following paragraphs are a sketch only of the development of the theory, in which an effort will be made to point out the memoirs which have been especially significant in the preparation of this paper. For more detailed references one should consult the articles on the calculus of variations in the *Encyclopädie der Mathematischen Wissenschaften* by Kneser [1, II A 8] * and Zermelo and Hahn [1, II A 8 a], the translations and extensions of them by Lecat in the *Encyclopédie des Sciences Mathématiques* [2], and the treatise by Bolza [3].

* The numbers in square brackets refer to the following bibliography.

Euler [2, p. 119; 7, p. 114] and Lagrange [8, I, p. 347] both studied special cases of the Lagrange problem which led up to the formulation of the more general problem and its multiplier rule by Lagrange [8, X, p. 420]. The proof of the multiplier rule which Lagrange gave was incomplete. The missing details were provided by A. Mayer [9], Hilbert [10], and Kneser [11, Sections 57-8]. Hahn [12] extended to the multiplier rule for the problem of Mayer, which includes that of Lagrange as a special case, the methods which Du Bois Reymond had applied to simpler problems of the calculus of variations. The argument in the text above is new but was suggested by papers by Hahn [13, p. 271] and Bliss [16].

The distinction between normal and abnormal minimizing arcs seems to have been first mentioned by A. Mayer [9, p. 79] but was emphasized by von Escherich [17] in connection with his theory of the second variation where it played an important role. Hahn [18, p. 152] adopts the definition of von Escherich. The definitions in Sections 7 and 8 above are modeled after that of Bolza [19, p. 440] and are applied to simplify the proof of the multiplier rule in Section 15 for the case when the functions ϕ_a contain no derivatives.

The necessary condition analogous to that of Legendre for simpler problems was first proved for the problem of Lagrange by Clebsch [20] as one of the consequences of his rather elaborate theory of the second variation. The necessary condition analogous to that of Weierstrass seems to have been first proved by Hahn [21] who deduced therefrom the necessary condition of Clebsch without appeal to the theory of the second variation. The method in the text above is that of Bolza [22], who supplied a step missing in the proof of Hahn, but the method is here further simplified by the use of the auxiliary formulas of Section 21 which are generalizations of formulas emphasized by Goursat [23, p. 566].

For the Lagrange problem the necessary condition for a minimum analogous to that of Jacobi for simpler problems is due to A. Mayer [24]. The envelope theorem and the associated geometric proof of the Mayer condition are the work of Kneser [25]. The method of the preceding pages for the development of Kneser's theory is modeled after Bolza [26], but with simplifications due again to the use of the auxiliary formulas of Section 21. The analytic proof of the Mayer condition by means of the theory of the minimum problem of the second variation was suggested by Bliss [27] and applied to the Lagrange problem by D. M. Smith [28]. By this method the advantages of the analytic proof are preserved without the necessity of using any complicated theory of the transformation of the second variation.

The theory of the second variation has been elaborately developed by many writers. The most important of the early papers is that of Clebsch [29] in which he transformed the second variation into its so-called reduced form and derived therefrom his necessary condition analogous to that of Legendre for simpler problems. The methods of Clebsch were modified by A. Mayer [30] who proved the necessity of a condition analogous to that of Jacobi for simpler problems, the so-called condition of Mayer described in the preceding pages. In a series of papers von Escherich [31] discussed in great detail the theory of the second variation and the various consequences which can be deduced from it. A condensed treatment of his theory is given by Bolza [32]. Hahn [33] showed the relationship between the theory of the second variation and certain aspects of the theories of Weierstrass as extended to the problem of Lagrange. The theory of the second variation takes a relatively simple form when it is viewed from the stand-point of the theory of the minimum problem of the second variation, as has been shown by Bliss [27, 34, 35].

The best reference for the sufficiency theorems in Chapter IV above is Bolza [36] to whom the precise formulation of the theorems and many details of the proofs are due. The properties of fields and their relation to the invariant integral analogous to that of Hilbert for simpler cases were first discussed by A. Mayer [37], and further material pertinent to the sufficiency proofs was discussed by Bolza [38] and Carathéodory [39]. The reader may refer to Kneser [11, 2d ed., pp. 290 ff.] for sufficiency proofs for the Mayer problem, and to Bliss [35] for a proof of the integral formula of Weierstrass and other properties of fields for the Lagrange problem.

BIBLIOGRAPHY.

GENERAL REFERENCES.

1. Kneser, "Variationsrechnung," *Encyclopädie der Mathematischen Wissenschaften*, II A 8; Zermelo und Hahn, "Weiterentwicklungen der Variationsrechnung in den letzten Jahren," *ibid.*, II A 8a.

2. Lecat, "Calcul des Variations," *Encyclopédie des Sciences Mathématiques*, II 31 (1913 and 1916).

3. Bolza, *Vorlesungen über Variationsrechnung* (1909).

4. Hadamard, *Leçons sur le Calcul des Variations* (1910).

5. Bliss, *The Calculus of Variations* (1924).

6. Moigno-Lindelöf, *Calcul des Variations* (1861).

THE EULER-LAGRANGE MULTIPLIER RULE.

7. Euler, *Methodus Inveniendi Lineas Curvas Maximi Minimive Proprietate Gaudentes* (1744).

8. Lagrange, *Oeuvres* I, X. See also the translation in Ostwald's *Klassiker der Exakten Wissenschaften*, Nr. 47, pp. 31 ff.

9. A. Mayer, "Begründung der Lagrangeschen Multiplicatorenmethode der Variationsrechnung," *Mathematische Annalen*, Vol. 26 (1886), pp. 74-82.

10. Hilbert, "Zur Variationsrechnung," *Göttinger Nachrichten* (1905), pp. 159-180; *Mathematische Annalen*, Vol. 62 (1906), pp. 351-368.

11. Kneser, *Lehrbuch der Variationsrechnung* (1900); 2d ed. (1925).

12. Hahn, "Über die Lagrangeschen Multiplicatorenmethode in der Variationsrechnung," *Monatshefte für Mathematik und Physik*, Vol. 14 (1908), pp. 325-342.

13. Hahn, "Über die Herleitung der Differentialgleichungen der Variationsrechnung," *Mathematische Annalen*, Vol. 63 (1907), pp. 253-272.

14. Bliss, "The Solutions of Differential Equations of the First Order as Functions of Their Initial Values," *Annals of Mathematics*, 2d Series, Vol. 6 (1905), pp. 49-68.

15. Bliss, "Solutions of Differential Equations as Functions of the Constants of Integration," *Bulletin of the American Mathematical Society*, Vol. 24 (1918), pp. 15-26.

16. Bliss, "The Problem of Mayer with Variable End-Points," *Transactions of the American Mathematical Society*, Vol. 19 (1918), pp. 305-314.

NORMAL ARCS.

17. von Escherich, see reference 31 below, Vol. 108 (1899), p. 1290.

18. Hahn, "Bemerkungen zur Variationsrechnung," *Mathematische Annalen*, Vol. 58 (1904), pp. 148-168.

19. Bolza, "Über den 'Anormalen Fall' beim Lagrangeschen und Mayerschen Problem mit gemischter Bedingungen und variabeln Endpunkten," *Mathematische Annalen*, Vol. 74 (1913), pp. 430-436.

THE NECESSARY CONDITIONS OF WEIERSTRASS AND CLEBSCH.

20. Clebsch, "Über die Reduction der zweiten Variation auf ihre einfachste Form," *Journal für die reine und angewandte Mathematik*, Vol. 55 (1858), pp. 254-273.

21. Hahn, "Über das allgemeine Problem der Variationsrechnung," *Monatshefte für Mathematik und Physik*, Vol. 17 (1906), pp. 295-304.

22. Bolza, see reference 3 above, pp. 603 ff.

23. Goursat, *Cours d'analyse*, Vol. 3, 4th ed. (1927), pp. 545-660.

THE NECESSARY CONDITION OF MAYER.

24. A. Mayer, "Über die Kriterien des Maximums und Minimums der einfachen Integrale," *Journal für die reine und angewandte Mathematik*, Vol. 69 (1868), pp. 238-263.

25. Kneser, "Die Jacobische Bedingung des Extremums bei einem allgemeinen Typus von Aufgaben der Variationsrechnung," *Mittheilungen der mathematischen Gesselschaft zu Kharkov*, Ser. 2, Vol. 7 (1902), pp. 253-267.

26. Bolza, see reference 3 above, pp. 613 ff.

27. Bliss, "Jacobi's Condition for Problems of the Calculus of Variations in Parametric Form," *Transactions of the American Mathematical Society*, Vol. 17 (1916), pp. 195-206.

28. D. M. Smith, "Jacobi's Condition for the Problem of Lagrange in the Calculus of Variations," *Transactions of the American Mathematical Society*, Vol. 17 (1916), pp. 459-475.

THE THEORY OF THE SECOND VARIATION.

29. Clebsch, see reference 20 above.

30. A. Mayer, see reference 24 above.

31. von Escherich, "Die zweite Variation der einfachen Integrale," *Sitzungsberichten der kaiserlichen Akademie der Wissenschaften in Wien*, Vol. 107 (1898), pp. 1191-1250, 1267-1326, 1383-1430; Vol. 108 (1899), pp. 1269-1340; Vol. 110 (1901), pp. 1355-1421.

32. Bolza, see reference 3 above, pp. 619-634.

33. Hahn, "Über den Zusammenhang zwischen den Theorien der zweiten Variation und der Weierstrass'schen Theorie die Variationsrechnung," *Rendiconti del Circolo Matematico di Palermo*, Vol. 29 (1910), pp. 49-78.

34. Bliss, "A Note on the Problem of Lagrange in the Calculus of Variations," *Bulletin of the American Mathematical Society*, Vol. 22 (1916), pp. 220-225.

35. Bliss, "The Transformation of Clebsch in the Calculus of Variations," *Proceedings of the International Mathematical Congress Held in Toronto*, Vol. 1 (1924), pp. 589-603.

SUFFICIENT CONDITIONS FOR A MINIMUM.

36. Bolza, see reference 3, pp. 635 ff.

37. A. Mayer, "Über den Hilbertschen Unabhangigkeitssatz in der Theorie des Maximums und Minimums der einfacher Integrale," *Leipziger Berichte*, Vol. 55 (1903), pp. 131-145.

38. Bolza, "Weierstrass' Theorem and Kneser's Theorem on Transversals for the Most General Case of an Extremum of simple Definite Integrals," *Transactions of the American Mathematical Society*, Vol. 7 (1906), pp. 459-488.

39. Carathéodory, "Die Methode der Geodätische Equidistanten und das Problem von Lagrange," *Acta Mathematica*, Vol. 47 (1926), pp. 199-236.

40. Kneser, see the reference 11, 2d ed., pp. 290 ff.

41. Bliss, see reference 35, p. 593.

Finite Geometries and the Theory of Groups.*

By R. D. CARMICHAEL.

Introduction.

The general purpose of this memoir is to exhibit the close contact which exists between the finite projective geometries $PG(k, p^n)$ and the theory of finite groups and to utilize the geometry in constructing permutation groups and in investigating their properties. Special attention is given to the case of multiply transitive permutation groups.

In the first division (§§ 1-6) a representation is given of the finite projective geometries $PG(k, p^n)$ by means of Abelian groups of type $(1, 1, 1, \cdots)$ and order $p^{(k+1)n}$ where p is prime. For the purpose of effecting this representation a system of coördinates for denoting the elements of such an Abelian group is introduced by means of the marks of the Galois field $GF[p^n]$. It is believed that these coördinates will be found useful for other purposes than those to which they are here put. They are used to aid in the selection and definition of a normal set of subgroups, which subgroups are interpreted as the points of a finite projective geometry $PG(k, p^n)$ of k dimensions. The elements themselves of the given Abelian group then become the points of a Euclidean geometry $EG(k+1, p^n)$ of $k+1$ dimensions. The theory of the finite geometries thus becomes available for developing the theory of Abelian groups of type $(1, 1, 1, \cdots)$, and *vice versa*. In particular, it is shown in § 6 that every theorem relating to a general projective space or a proper projective space or a modular projective space or a rational modular projective space (in the sense of Veblen and Young, l. c.) may be translated into a theorem about Abelian groups of type $(1, 1, 1, \cdots)$. Thus by a single act of thought a significant extension is given to the theory of Abelian groups and a method is made apparent by which the theory may be further developed.

By means of the coördinates introduced in § 2 to denote the elements of an Abelian group G of order $p^{(k+1)n}$ and of type $(1, 1, 1, \cdots)$ analytical representations are set up in the second division of the memoir (§§ 7-11) for the group of isomorphisms of the named Abelian group and for its holomorph. These representations afford generalisations of known results. Incidentally to the study of certain subgroups of the group of isomorphisms a generalisation of the Betti-Mathieu group appears (§ 9). Finally, in the

* Presented to the American Mathematical Society (Kansas City), Dec. 29, 1925.

last section of this division certain transformation groups in $PG(k+1, p^n)$ are formed from the earlier groups in the division by aid of the interpretation of the given Abelian group by means of the Euclidean space $EG(k+1, p^n)$.

The third division of the memoir (§§ 12-14) is devoted to the development of certain central theorems concerning collineation groups in the finite geometries and their subgroups. The main results are given in the first and third theorems of § 12 and the first theorem of § 13. The results include and generalise several known theorems concerning doubly transitive and triply transitive groups. In particular the existence is shown of several infinite classes of triply transitive and of doubly transitive groups, including certain such classes already known. Moreover, infinite classes of simply transitive primitive groups are also exhibited. It is proved that there is no upper limit K to the number of primitive groups (of varying degrees) in a set of primitive groups each of which is simply isomorphic with each of the others in the set. Furthermore, it is shown that, for every integer L there exist integers $s[t]$ such that the number of the doubly transitive [triply transitive] groups of degree $s[t]$ is greater than L.

I. Representation of the Finite Projective Geometries $PG(k, p^n)$ by Means of Abelian Groups.

1. *The Finite Projective Geometry $PG(k, p^n)$.* Let p be any prime number and k be any positive integer. Let us consider the Abelian group G_{k+1} of order p^{k+1} and type $(1, 1, 1, \cdots)$. Every element of this group except the identity is of order p. The number of these elements is $p^{k+1}-1$. Each of them generates a subgroup of order p, the same subgroup being generated by each of $p-1$ different elements. Hence the group G_{k+1} contains

$$(p^{k+1}-1)/(p-1), \text{ or } 1+p+p^2+\cdots+p^k,$$

distinct subgroups of order p. The totality of these subgroups contains all the elements of G_{k+1}; and no two of these subgroups have any element in common except identity.

Each of these subgroups of order p in G_{k+1} will be called a point in the finite geometry $PG(k, p)$ which we are engaged in constructing. This k-dimensional finite geometry then contains just $1+p+p^2+\cdots+p^k$ points.

Now consider any two points of the $PG(k, p)$. From the group-theoretic point of view they are two subgroups of G_{k+1} of order p. The group generated by them is of order p^2 and type $(1, 1)$. It contains $1+p$ subgroups of order p; and no two of these subgroups have any element in common except

identity. From the geometric point of view these $1+p$ subgroups are $1+p$ points of the $PG(k,p)$. We shall say that they form a line in the $PG(k,p)$. Thus any two points in $PG(k,p)$ determine a line of $PG(k,p)$, and this line has just $1+p$ points on it. We shall denote by AB the line containing the two distinct points A and B.

Let us determine the number of lines in $PG(k,p)$. In determining a line we may select a first point in $1+p+p^2+\cdots+p^k$ ways and then a second point in $p+p^2+\cdots+p^k$ ways. But this procedure will select the same line in as many ways as two points on it may be chosen in an assigned order. The first point may be taken in $1+p$ ways and then the second in p ways. Hence the number of lines in $PG(k,p)$ is

$$(1+p+p^2+\cdots+p^k)(p+p^2+\cdots p^k)/(1+p)p$$

or

$$(p^{k+1}-1)(p^k-1)/(p^2-1)(p-1).$$

This of course is the same as the number of subgroups of order p^2 in G_{k+1}.

An m-dimensional space in $PG(k,p)$, $m \leqq k$, may now be defined as the set of points each of which is identified with the corresponding subgroup of order p in a given subgroup of G_{k+1} of order p^{m+1}, its type of course being necessarily $(1,1,1,\cdots)$. For $m=2$ we have the case of a plane. The number of points in the m-dimensional space is

$$1+p+p^2+\cdots+p^m.$$

It is obvious that this m-dimensional space is completely determined by any $m+1$ of its points so selected that they do not all lie in any $(m-1)$-dimensional space.

In this m-dimensional space $PG(m,p)$ there are included $(m-1)$-dimensional spaces $PG(m-1,p)$. Let us consider any such space S_{m-1} of $m-1$ dimensions, m now being greater than 1; and let P be a point not in S_{m-1}. Let T be the set of points each of which is collinear (on the same line) with P and some point of S_{m-1}. From the group-theoretic interpretation it is clear that the set of points T constitute an m-dimensional space $PG(m,p)$. Thus we may have an inductive definition of the points of a space of m dimensions. A point is a 0-space. If $P_1, P_2, \cdots, P_{m+1}$ are points not all in the same $(m-1)$-space, then the set of all points each of which is collinear with P_{m+1} and some point of the $(m-1)$-space $(P_1, P_2, \cdots, P_m)$ is the m-space $(P_1, P_2, \cdots, P_{m+1})$. It is obvious that this inductive definition is equivalent to the definition already given.

The number of ways in which $m+1$ points may be selected in a given order so that they do not all lie in any $(m-1)$-dimensional space is

$$(1+p+p^2+\cdots+p^k)(p+p^2+\cdots p^k)$$
$$\times(p^2+p^3+\cdots+p^k)\cdots(p^m+p^{m+1}+\cdots+p^k),$$

the factors of this expression in the order written being the number of ways in which the first point, the second point, $\cdots$, the $(m+1)$-th point, respectively, may be selected. The number of ways in which $m+1$ points of a given $PG(m,p)$ may be selected in a given order so that they do not all lie on any $(m-1)$-dimensional space is

$$(1+p+\cdots+p^m)(p+p^2+\cdots+p^m)$$
$$\times(p^2+p^3+p^m)\cdots(p^{m-1}+p^m)p^m.$$

It is obvious that the number of m-dimensional spaces $PG(m,p)$ in the given $PG(k,p)$ is the quotient of the first of the two foregoing products divided by the second; this quotient may be written in the form

$$\frac{(p^{k+1}-1)(p^k-1)(p^{k-1}-1)\cdots(p^{k-m+1}-1)}{(p^{m+1}-1)(p^m-1)(p^{m-1}-1)\cdots(p^2-1)(p-1)}.$$

This of course is the same as the number of subgroups in G_{k+1} of order p^{m+1}.

When $m+1$ generators of G_{k+1} are selected for generating the subgroups corresponding to a given $PG(m,p)$ there are left in G_{k+1} $k-m$ other independent generators independent of the $m+1$ already employed. These give rise to a $PG(k-m-1,p)$. Thence we see that the number of m-spaces in $PG(k,p)$ is the same as the number of $(k-m-1)$-spaces. In particular, the number of points in a plane is equal to the number of lines in the plane.

Veblen and Bussey * define a finite projective geometry in the following way. It consists of a set of elements, called points for suggestiveness, which are subject to the following five conditions or postulates:

I. The set contains a finite number (>2) of points. It contains one or more subsets called lines, each of which contains at least three points.

II. If A and B are distinct points, there is one and only one line that contains both A and B.

III. If A, B, C are non-collinear points and if a line l contains a point D of the line AB and a point E of the line BC but does not contain A or B or C, then the line l contains a point F of the line CA.

IV_k. If m is an integer less than k, not all of the points considered are in the same m-space.

V_k. If IV_k is satisfied, there exists in the set of points considered no $(k+1)$-space.

* O. Veblen and W. H. Bussey, "Finite Projective Geometrics," *Transactions of the American Mathematical Society*, Vol. 7 (1906), pp. 241-259.

The geometry so defined is a geometry of k-dimensional space.

In this system of postulates the terms point and line are left undefined. A point is called a 0-space and a line is called a 1-space. Spaces of higher dimensions are defined inductively by the method which we have already shown to be equivalent to our first definition of an m-space. To show that the set of points which we have defined constitute a finite projective geometry it is therefore sufficient to prove that each of the foregoing postulates is satisfied. From the properties of the group G_{k+1} it follows at once that postulates I, II, IV_k, V_k are satisfied by the set of points in $PG(k, p)$. It remains to show that postulate III is verified. For this purpose let a, b, c be generators of the subgroups of G_{k+1} corresponding to the points A, B, C respectively. Then the groups corresponding to the points of the lines AB, BC, CA have respectively as generators the elements

$$a^{\alpha_1}b^{\beta_1}, \quad b^{\beta_2}c^{\alpha_2}, \quad c^{\gamma_3}a^{\alpha_3},$$

where each exponent belongs to the set $0, 1, 2, \cdots, p-1$ and at least one exponent in the symbol for each generator is different from zero. If a generator of the group corresponding to D in the postulate is $a^{\alpha}b^{\beta}$ then both α and β belong to the set $1, 2, \cdots, p-1$, since D is different from A and B. Likewise both ρ and σ in a generator $b^{\rho}c^{\sigma}$ of the group corresponding to E belong to the set $1, 2, \cdots, p-1$. Then the line DE corresponds to the group $\{a^{\alpha}b^{\beta}, b^{\rho}c^{\sigma}\}$. The elements in this group are $a^{\lambda\alpha}\, b^{\lambda\beta}\, b^{\mu\rho}\, c^{\mu\sigma}$ where λ and μ range independently over the set $0, 1, 2, \cdots, p-1$. Now λ and μ, both different form zero, exist such that $\lambda\beta + \mu\rho \equiv 0$ modulo p. The corresponding element of the group is then $a^{\lambda\alpha}c^{\mu\sigma}$. This generates a group corresponding to a point on the line AC; it is different from A and C since each of the numbers α, σ, λ, μ is incongruent to zero modulo p. This is the point F common to DE and CA whose existence is asserted by postulate III. Hence the set of points in our $PG(k, p)$ satisfies the foregoing postulates and therefore constitutes a finite projective geometry.*

It is desirable to introduce homogeneous coördinates for representing the points in the finite projective geometry $PG(k, p)$. For this purpose let us consider a set of $k+1$ independent generators $a_0, a_1, a_2, \cdots, a_k$ of the group G_{k+1}. Then the elements of this group are all represented uniquely by the set of symbols

$$a_0^{\mu_0}a_1^{\mu_1}a_2^{\mu_2}\cdots a_k^{\mu_k}$$

* The special case of the geometry $PG(3, 2)$ is treated briefly in a manner similar to the foregoing by U. G. Mitchell in his dissertation (footnote on p. 34).

where $\mu_0, \mu_1, \cdots, \mu_k$ run independently over the set $0, 1, 2, \cdots, p-1$ of p numbers. An element of G_{k+1} may therefore be denoted uniquely by the symbol

$$\{\mu_0, \mu_1, \mu_2, \cdots, \mu_k\}$$

where each μ is a number of the set $0, 1, 2, \cdots, p-1$, provided it is understood that the symbol represents the product $a_0^{\mu_0}a_1^{\mu_1}\cdots a_k^{\mu_k}$, the a's forming a fixed set of independent generators of G_{k+1}. Two such symbols are to be considered equivalent if their corresponding elements are congruent modulo p. For the multiplication of these symbols (corresponding to multiplication of elements in G_{k+1}) we obviously have the following formula

$$\{\mu_0, \mu_1, \cdots, \mu_k\}\{\nu_0, \nu_1, \cdots, \nu_k\} = \{\mu_0 + \nu_0, \mu_1 + \nu_1, \cdots, \mu_k + \nu_k\}.$$

Now consider the set of elements

$$\{\mu\mu_0, \mu\mu_1, \cdots, \mu\mu_k\}$$

where $\mu_0, \mu_1, \cdots, \mu_k$ constitute a fixed set of $k+1$ numbers taken modulo p and not all of them are congruent to zero modulo p, μ being a variable integer taken modulo p. It is easy to see that this set of elements forms a group of order p having $\{\mu_0, \mu_1, \cdots, \mu_k\}$ for a generator. This group may be denoted by the symbol

$$(\mu_0, \mu_1, \cdots, \mu_k).$$

The same group is also represented by the symbol

$$(\rho\mu_0, \rho\mu_1, \cdots, \rho\mu_k)$$

provided only that ρ is a fixed integer incongruent to zero modulo p. The corresponding point will be denoted by the symbol

$$(\mu_0, \mu_1, \cdots, \mu_k)$$

and $\mu_0, \mu_1, \cdots, \mu_k$ will be called homogeneous coördinates of the point. The condition that such a symbol shall represent a point is that the μ's shall be integers and that one of them at least shall be different from zero modulo p. Two such symbols represent the same point if the corresponding coördinates are proportional modulo p. Except for this factor of proportionality there is thus a unique correspondence between the points of $PG(k, p)$ and the symbols which represent them by means of coördinates.

2. *Generalization to the Finite Projective Geometry $PG(k, p^n)$.* Let us consider more generally an Abelian group $G_{(k+1)n}$ of order $p^{(k+1)n}$ and type $(1, 1, 1, \cdots)$, p being a prime number and k and n being any positive in-

tegers. The points of our finite geometry $PG(k, p^n)$ are to be certain subgroups of $G_{(k+1)n}$ of order p^n. To begin with, these subgroups * are to be selected in such a way † that no two of them shall have any element in common except identity and so that the set shall contain all the elements of $G_{(k+1)n}$. The number of elements other than identity in $G_{(k+1)n}$ is $p^{(k+1)n} - 1$; and the number of such elements in a subgroup of order p^n is $p^n - 1$. Hence a set of subgroups of $G_{(k+1)n}$ of order p^n and having the properties named will consist of

$$(p^{(k+1)n} - 1)/(p^n - 1), \text{ or } 1 + p^n + p^{2n} + \cdots + p^{kn},$$

subgroups. Therefore the k-dimensional geometry $PG(k, p^n)$, to be defined, will consist of $1 + p^n + p^{2n} + \cdots + p^{kn}$ points.

In order to select an appropriate set of subgroups of order p^n for the purpose in hand we shall first develop a method of representing the elements of $G_{(k+1)n}$ by means of the marks of a Galois field, thus generalizing the results at the end of the preceding section. This mode of representing the elements of an Abelian group of type $(1, 1, 1, \cdots)$ we shall find useful for other purposes besides the geometrical one which now engages our attention.

Let us denote a set of $(k+1)n$ independent generating elements of $G_{(k+1)n}$ by

$$\begin{array}{ccccc} a_{01}, & a_{02}, & a_{03}, \cdots, & a_{0n}, \\ a_{11}, & a_{12}, & a_{13}, \cdots, & a_{1n}, \\ \cdot & \cdot & \cdot & \cdot \\ a_{k1}, & a_{k2}, & a_{k3}, \cdots, & a_{kn}. \end{array}$$

Then every element in $G_{(k+1)n}$ may be represented uniquely in the form

$$\prod_{i=0}^{k} a_{i1}^{s_{i1}} a_{i2}^{s_{i2}} \cdots a_{in}^{s_{in}}$$

where the exponents s are integers taken modulo p. The element denoted by this product for a fixed set of exponents s will be represented by the symbol

$$\{\mu_0, \mu_1, \mu_2, \cdots, \mu_k\}$$

where μ_i $(i = 0, 1, 2, \cdots, k)$ denotes that mark of the Galois field $GF[p^n]$ which may be written in the form

$$\mu_i = s_{i1} + s_{i2}\omega + s_{i3}\omega^2 + \cdots + s_{in}\omega^{n-1},$$

* The special case when $p = 2$ and $k = 1$ is treated incidentally (in a different manner) by L. E. Dickson, *Bulletin of the American Mathematical Society*, Ser. 2, Vol. 11 (1905), pp. 177-179.

† See G. A. Miller, *Bulletin of the American Mathematical Society*, Ser. 2, Vol. 12 (1906), pp. 446-449, for theorems relating to this problem.

ω being a fixed primitive mark of $GF[p^n]$. This correspondence of elements and symbols is unique in the sense that to each element there corresponds a single symbol and to each symbol there corresponds a single element.

For the multiplication of these symbols, corresponding to the multiplication of elements in $G_{(k+1)n}$, we have the following obvious formula:

$$\{\mu_0, \mu_1, \cdots, \mu_k\}\{\nu_0, \nu_1, \cdots, \nu_k\} = \{\mu_0 + \nu_0, \mu_1 + \nu_1, \cdots, \mu_k + \nu_k\}.$$

Now suppose that $\mu_0, \mu_1, \cdots, \mu_k$ is a fixed set of $k + 1$ marks of $GF[p^n]$, at least one of them being different from zero; and consider the set of elements

$$\{\mu\mu_0, \mu\mu_1, \cdots, \mu\mu_k\}$$

where μ is a variable running over the p^n marks of $GF[p^n]$. It is obvious that the elements in this set are all distinct and that their number is p^n. Moreover, the product of any two of them is in the set, as one sees immediately from the law of multiplication and the properties of the marks of a Galois field. This set of elements therefore constitutes a subgroup of $G_{(k+1)n}$ of order p^n. It is easy to see that the elements

$$\{\omega^i\mu_0, \omega^i\mu_1, \cdots, \omega^i\mu_k\}, \qquad (i = 0, 1, 2, \cdots, n-1),$$

constitute a set of independent generators of this subgroup. If σ is any non-zero mark of the Galois field the same subgroup obviously consists of the set of elements

$$\{\mu\sigma\mu_0, \mu\sigma\mu_1, \cdots, \mu\sigma\mu_k\},$$

μ varifying as before. The subgroup itself may therefore be represented by the symbol

$$(\mu_0, \mu_1, \cdots, \mu_k)$$

where $\mu_0, \mu_1, \cdots, \mu_k$ are interpreted as the "homogeneous coördinates" of the subgroup. On multiplying each of the coördinates by one and the same non-zero mark of the field we have merely proportional homogeneous coördinates of the same subgroup. To each set of ordered coördinates, one at least of the coördinates being different from zero, there corresponds a subgroup of $G_{(k+1)n}$ of order p^n.

The number of subgroups in the set denoted by $(\mu_0, \mu_1, \cdots, \mu_k)$ for varying μ's is readily determined. Each symbol μ may be chosen in p^n independent ways except that they cannot all be zero. Hence the number of choices is $p^{(k+1)n} - 1$. To obtain the number of subgroups we must divide

this by the number $p^n - 1$ of possible factors of proportionality in the various notations for the same group. Hence the number of groups in our set is

$$(p^{(k+1)n} - 1)/(p^n - 1), \text{ or } 1 + p^n + p^{2n} + \cdots + p^{kn}.$$

Once $G_{(k+1)n}$ has been given, this selection of groups depends on two things: the ordered set of $(k+1)n$ independent generators and the primitive mark ω by means of which the marks μ_i were first introduced. With reference to this selected basis of determination we shall call the set of subgroups just determined a normal set. By means of other sets of generators and other primitive marks we might in certain cases select other normal sets of subgroups of $G_{(k+1)n}$. Since we shall use the same basis throughout this memoir we shall speak of the foregoing normal set of subgroups without reference to the basis on which it has been defined.

For the case $n = 1$, it is to be observed, a subgroup of a normal set is simply any subgroup of order p.

No two subgroups of a normal set have any element in common except identity, as one may readily prove by means of the symbols which represent their elements. Moreover a given element of $G_{(k+1)n}$ occurs in some subgroup of a normal set. Hence the subgroups of a normal set have the properties demanded at the beginning of the section for points. Accordingly for the points of $PG(k, p^n)$ we take the subgroups of a normal set of subgroups of $G_{(k+1)n}$. That the latter group has (when $n > 1$) other subgroups of the same order p^n will not concern us at the present.

We shall represent a point of $PG(k, p^n)$ by the same symbol

$$(\mu_0, \mu_1, \cdots, \mu_k)$$

as we have already employed to denote the subgroup of order p^n which we identify with this point. Thus we have a set of homogeneous coördinates to represent the points of $PG(k, p^n)$, each one of the coördinates being a mark of $GF[p^n]$.

An m-dimensional space, or an m-space, in $PG(k, p^n)$, $m \leqq k$, may now be defined as the set of points corresponding to the groups of a normal set of subgroups of $G_{(k+1)n}$ which are contained as subgroups in the group generated by $m+1$ of the groups of a normal set, these $m+1$ groups being such that no one of them is contained in the group generated by the other m. A point will be called a 0-space; a 1-space will be called a line; a 2-space we will call a plane. It is clear that this definition is again equivalent to the inductive definition given in § 1 for the special case when $n = 1$.

To show that the $PG(k, p^n)$ is a finite projective geometry in the sense

6

of Veblen and Bussey we have now only to prove that the postulates given in § 1 are verified when interpreted as referring to our $PG(k, p^n)$. That postulates I, II, IV_k, V_k hold is immediately obvious. It remains only to verify postulate III.

For this purpose consider three non-collinear points A, B, C and let

$$(\alpha_0, \alpha_1, \cdots, \alpha_k), \quad (\beta_0, \beta_1, \cdots, \beta_k), \quad (\gamma_0, \gamma_1, \cdots, \gamma_k)$$

respectively be their coördinates. Then a point D on the line AB determined by the points A and B has the coördinates

$$(\alpha\alpha_0 + \beta\beta_0, \alpha\alpha_1 + \beta\beta_1, \cdots, \alpha\alpha_k + \beta\beta_k)$$

where α and β are marks of $GF[p^n]$. A necessary and sufficient condition that this point shall be different from A and B is that both α and β shall be different from zero. Hence we take them to be different from zero. Likewise a point E on BC has the coördinates

$$(\rho\beta_0 + \sigma\gamma_0, \rho\beta_1 + \sigma\gamma_1, \cdots, \rho\beta_k + \sigma\gamma_k)$$

where ρ and σ are marks of $GF[p^n]$. We take ρ and σ to be both different from zero so that E shall be different from both B and C. Now a point on the line DE has the coördinates

$$(\lambda\alpha\alpha_0 + \lambda\beta\beta_0 + \mu\rho\beta_0 + \mu\sigma\gamma_0, \cdots, \lambda\alpha\alpha_k + \lambda\beta\beta_k + \mu\rho\beta_k + \mu\sigma\gamma_k)$$

where λ and μ are marks of $GF[p^n]$. Since β and ρ are both different from zero there exist non-zero marks λ and μ such that $\lambda\beta + \mu\rho$ is zero. For such a pair of values of λ and μ the corresponding point F of DE has the coördinates

$$(\lambda\alpha\alpha_0 + \mu\sigma\gamma_0, \cdots, \lambda\alpha\alpha_k + \mu\sigma\gamma_k).$$

This F is a point on the line CA; and it is different from both C and A, since each of the marks α, β, λ, μ is different from zero. From the relation thus established among the points A, B, C, D, E, F it is seen that postulate III is verified.

Hence, the $PG(k, p^n)$, as we have defined it, is a finite projective geometry in the sense of Veblen and Bussey.

Veblen and Bussey (*loc. cit.*) proved that when $k > 2$ every finite projective k-dimensional geometry satisfying the definition which we have reproduced in § 1 is a geometry of points whose homogeneous coördinates may be taken as the marks of $GF[p^n]$ in precisely the same way as we have used homogeneous coördinates to represent the points of our $PG(k, p^n)$. This justifies us in using for these geometries the symbol $PG(k, p^n)$ already em-

ployed by Veblen and Bussey. Moreover, we may say that the foregoing group-theoretic construction of $PG(k, p^n)$ affords an interpretation in the theory of Abelian groups of type $(1, 1, 1, \cdots)$ of every possible finite projective geometry of more than two dimensions. We shall not now treat the problem of possible group-theoretic interpretations of the remaining finite geometries, namely, certain of those of two dimensions.

It is now evident that every theorem relating to $PG(k, p^n)$ can be translated into a corresponding theorem about the group $G_{(k+1)n}$. We shall illustrate the remark by so interpreting the following geometric theorem:

If l and m are positive integers less than k and such that $l + m - k = r \geqq 0$, then, in the given k-space, an l-space and an m-space have at least an r-space in common.

The group-theoretic interpretation is as follows:

Let $s_1, s_2, \cdots, s_{l+1}$ be any $l + 1$ subgroups of a normal set of subgroups of $G_{(k+1)n}$ such that no one of them is contained in the group generated by the other l, and let $\sigma_1, \sigma_2, \cdots, \sigma_{m+1}$ be a like set of $m + 1$ such subgroups. Then, if $l < k$, $m < k$, $l + m - k = r \geqq 0$, the groups $\{s_1, s_2, \cdots, s_{l+1}\}$ and $\{\sigma_1, \sigma_2, \cdots, \sigma_{m+1}\}$ contain at least $r + 1$ subgroups of a normal set such that no one of these subgroups is contained in the group generated by the remaining r of them.

The number of m-spaces $PG(m, p^n)$, $m < k$, contained in the given k-space $PG(k, p^n)$ is readily determined in the general case by the same method as that employed in § 1 for the special case $n = 1$. This number turns out to be

$$\frac{(p^{(k+1)n} - 1)(p^{kn} - 1)(p^{(k-1)n} - 1) \cdots (p^{(k-m+1)n} - 1)}{(p^{(m+1)n} - 1)(p^{mn} - 1)(p^{(m-1)n} - 1) \cdots (p^n - 1)}.$$

In the foregoing part of the section we have given an analytic method for determining normal sets of subgroups of $G_{(k+1)n}$. It is desirable to have such a set characterised by means of properties which are immediately group-theoretic in their character. The subgroups of a given normal set have the following properties and mutual relations, as we have already seen:

1) Each of these subgroups is of order p^n.
2) No two of these subgroups have a common element except identity.
3) Any given element of $G_{(k+1)n}$ is contained in some subgroup of a normal set.
4) If A, B, C are three subgroups of a normal set such that no one of

them is in the group generated by the other two, and if D is a subgroup of the group $\{A, B\}$ and is different from A and B and belongs to the normal set, and finally if E is a subgroup of the group $\{B, C\}$ and is different from B and C and belongs to the normal set, then the groups $\{C, A\}$ and $\{D, E\}$ have in common a group F which belongs to the normal set.

Now any set of subgroups of $G_{(k+1)n}$ which have these properties alone clearly satisfy the five defining postulates given in § 1. They therefore afford a representation of a finite geometry. But Veblen and Bussey (*loc. cit.*) have shown that every finite projective k-dimensional geometry satisfying the definition reproduced in § 1 is a $PG(k, p^n)$, in the sense of their use of this symbol, provided that $k > 2$. Hence one can introduce coördinates into this geometry by means of the marks of a Galois field. On doing this in the case of the given group-theoretic representation of the geometry we exhibit the set of subgroups involved as a normal set in accordance with the definition of such a set. Therefore when $k > 2$ the properties 1), 2), 3), 4) of a normal set of subgroups furnish a complete group-theoretic characterization of such a set. The conclusion will also hold for $k = 1$ or 2 if we suppose that the normal set of subgroups is so chosen that it may be taken as a part of the normal set of subgroups in a group of order p^{4n} and type $(1, 1, 1, \cdots)$ which contains the given group $G_{(k+1)n}$ for $k = 1$ or 2.

Consider now the $PG(k+1, p^n)$ whose points are denoted by the symbols $(\mu_0, \mu_1, \cdots, \mu_{k+1})$ where each μ is a mark of $GF[p^n]$. Those points for which μ_{k+1} is different from zero constitute the Euclidean finite geometry $EG(k+1, p^n)$, this being obtained by omitting from $PG(k+1, p^n)$ those points for which the last coördinate is zero. For the points of $EG(k+1, p^n)$ we may take $\mu_{k+1} = 1$. Then the coördinates $\mu_0, \mu_1, \cdots, \mu_k$ may be taken as the non-homogeneous coördinates of points in $EG(k+1, p^n)$. Such a point $(\mu_0, \mu_1, \cdots, \mu_k, 1)$ may then be identified with the element $\{\mu_0, \mu_1, \cdots, \mu_k\}$ of $G_{(k+1)n}$. Hence the elements of this group may be taken as the points of the Euclidean finite geometry $EG(k+1, p^n)$. Hence the theorems in the latter geometry may be interpreted as theorems concerning the elements of $G_{(k+1)n}$.

3. *The Principle of Duality.* The principle of duality is valid in the finite geometry $PG(k, p^n)$. If l is less than k the dual of the set of l-spaces in $PG(k, p^n)$ is the set of $(k-l-1)$-spaces. In particular the dual of the set of points in $PG(k, p^n)$ is the set of $(k-1)$-spaces contained in the given k-space. Since the number of elements [l-spaces] in a set of subspaces is

equal to the number of elements $[(k-l-1)$-spaces] in the dual set of spaces, it follows in particular that the number of subgroups of a normal set of subgroups of $G_{(k+1)n}$ is equal to the number of subgroups each of which is generated by $k-1$ independent subgroups of the normal set. For $n=1$ this reduces to the well known theorem that the number of subgroups of order p in an Abelian group of order p^{k+1} and type $(1, 1, 1, \cdots)$ is equal to the number of subgroups of index p. More generally, if l is less than $k+1$ the number of subgroups of order p^l in this group G_{k+1} is equal to the number of subgroups of index p^l.

In general every theorem about the Abelian group of order p^m and type $(1, 1, 1, \cdots)$, which is capable of interpretation as a theorem in a finite geometry $PG(k, p^n)$, may be dualized. It will thus lead to a new theorem about the Abelian group, except in the special case when the theorem is its own dual. For the purpose of obtaining these theorems about a given Abelian group of order p^m and type $(1, 1, 1, \cdots)$, one may construct a corresponding geometry $PG(k, p^n)$ for every pair of positive integral values of k and n such that $(k+1)n=m$. Thus if m is highly composite the given Abelian group may be investigated by means of any one of several finite geometries constructed in the manner indicated. The case $k+1=m$ and $n=1$ will be especially useful for this purpose since in this case the normal set of subgroups consists of all the subgroups of order p.

From the principle of duality it follows that one of the requirements for points named at the beginning of § 2 is superfluous, at least in the form there stated. It was prescribed that the subgroups which were to represent points were to be selected in such a way that no two of them should have any element in common except identity. Now that the geometry has been constructed a new one can be made from it such that the points in the new geometry are the dual elements of the points in the old geometry. In this new geometry two given points, when considered as subgroups, will have elements in common besides the identity. And yet the new geometry will serve equally well as a means of investigating the given Abelian group.

Once this general principle of duality in the theory of Abelian groups is recognized, a number of properties of these groups heretofore discovered become almost or quite obvious, since a fundamental reason for their appearance is manifest.

4. *The Complete Quadrangle.* In the finite projective geometries $PG(k, p^n)$ there is an important distinction to be made according as the prime p is equal to 2 or is odd. This distinction was investigated by Veblen

and Bussey in the article cited. They showed (p. 245) that the diagonal points of a complete quadrangle are collinear when $p=2$ and are non-collinear when p is an odd prime. Thus an important and simple geometric fact sharply distinguishes between the two named cases of these finite geometries.

This difference in the geometries (for the two cases) must be reflected in an important way in the theory of Abelian groups of order p^m and type $(1, 1, 1, \cdots)$. Early in the development of the theory of these groups it was noticed that their properties differ owing to whether p is 2 or is an odd prime. From our geometric interpretation and the facts just stated (in this section), the fundamental basis for this difference is apparent. Hence, in investigating these groups, one will now know precisely from what place to begin to develop those features of the theory which depend on the even or odd character of p.

For the case of the Abelian group $G_{(k+1)n}$, with the geometry $PG(k, p^n)$ constructed from it, the distinguishing difference of the two cases may be stated in group-theory language as follows (it being assumed now that $k>1$): Let A, B, C, D be four subgroups of a normal set of subgroups of $G_{(k+1)n}$ such that no one of them is contained in the group generated by another two while D is contained in the group $\{A, B, C\}$. Let E be the (unique) subgroup of the normal set common to the groups $\{A, B\}$ and $\{C, D\}$, F that common to the groups $\{A, C\}$ and $\{B, D\}$ and G that common to the groups $\{A, D\}$ and $\{B, C\}$. Then each of the subgroups E, F, G is in the subgroup generated by the other two when and only when $p=2$.

A large part of the theory of the geometry $PG(k, p^n)$ can be developed independently of any hypothesis as to the collinearity or noncollinearity of the diagonal points of a complete quadrangle (see § 6 of this paper). These theorems will give rise to corresponding theorems about Abelian groups of order p^m and type $(1, 1, 1, \cdots)$ which are independent of the odd or even character of p.

5. *The Theorems of Desargues and Pascal.* As an example of another interesting theorem in the theory of groups obtained from a geometric fact, let us consider the following.

The theorem of Desargues, which is valid in the $PG(k, p^n)$, may be stated thus. Let ABC and abc be two triangles in the same plane and let them be perspective from a point O so that O, A, a are collinear, O, B, b are collinear, and O, C, c are collinear. Let γ be the point of intersection of AB and ab, β that of AC and ac, and α that of BC and bc. Then the points α, β, γ are collinear.

Let us translate this result into a theorem concerning the Abelian group

$G_{(k+1)n}$ viewed as indicated in § 2 in the light afforded by the geometry $PG(k, p^n)$, it being assumed now that $k > 1$.

Let A, B, C be three subgroups of a normal set of subgroups of $G_{(k+1)n}$ such that no one of them is in the group generated by the other two. We select other subgroups of the normal set as follows, each of them to be in the group $\{A, B, C\}$: O is any such subgroup which is not contained in any one of the groups $\{A, B\}$, $\{B, C\}$, $\{C, A\}$; a, b, c are such subgroups different from O, A, B, C and contained respectively in the groups $\{O, A\}$, $\{O, B\}$, $\{O, C\}$. Let γ, α, β be the subgroups of the normal set of sub-subgroups common to the respective pairs of groups

$$\{A, B\},\ \{a, b\};\quad \{B, C\},\ \{b, c\};\quad \{C, A\},\ \{c, a\}.$$

Then each of the subgroups α, β, γ is in the subgroup generated by the other two.

The generalizations of the theorem of Desargues to higher dimensions yield likewise interesting theorems concerning Abelian groups. As phrased abstractly the theorems seem to be rather complicated; but in their geometric formulation they are easily comprehended and retained in mind.

As affording a final illustration of this method of translating geometric theorems into theorems about Abelian groups, let us consider the following which gives rise to the configuration of Pappus (Veblen and Young, *Projective Geometry,* Vol. I, p. 98). If A, B, C are any three distinct points of a line l, and A', B', C', are any three additional distinct points on another line l' meeting l in O, the three points γ, α, β of intersection of the respective pairs of lines

$$AB',\ A'B;\quad BC',\ B'C;\quad CA',\ C'A$$

are collinear.

Translating as in the previous case we have the following theorem:

Let O, A, A' be three subgroups of a normal set of subgroups of $G_{(k+1)n}$ such that no one of them is in the group generated by the other two. Let B and C be two additional subgroups contained in the group $\{O, A\}$ and belonging to the normal set, and B' and C' be two additional such subgroups contained in the group $\{O, A'\}$, these groups being existent when and only when $p^n > 2$ and $k > 1$. Let γ, α, β be the subgroups of the normal set which are common to the respective pairs of groups

$$\{A, B'\}, \{A', B\};\quad \{B, C'\}, \{B', C\};\quad \{C, A'\}, \{C', A\}.$$

Then each of the subgroups α, β, γ is in the subgroup generated by the other two.

6. *Geometries Affording Applications to Abelian Groups.* The analysis and development of projective geometry given by O. Veblen and J. W. Young (*Projective Geometry,* Vol. I, 1910; Vol. II, 1918) afford a convenient means of ascertaining what geometries have direct applications to the theory of Abelian groups by means of the representations of finite geometries given in the foregoing pages. In vol. II (p. 36) of this work, Veblen describes nine classes of geometries characterized by means of the assumptions which underlie them. Using capital letters to denote the assumptions and employing the notation of Veblen and Young (see the index to vol. II under the word "Assumption"), we select for our purpose four of these geometries as follows: A space satisfying Assumptions

A, E	is a general projective space;
A, E, P	is a proper projective space;
$A, E, \bar{H}$	is a modular projective space;
$A, E, \bar{H}, Q$	is a rational modular projective space.

It is easy to verify that the assumptions involved in these four geometries are all valid in the case of the geometry $PG(k, p^n)$, except that Q is valid when and only when $n = 1$. Since the points of this geometry have been represented by certain subgroups of the Abelian group $G_{(k+1)n}$, it follows that every theorem in any one of the four geometries named is capable of immediate translation into a theorem concerning the given Abelian group. In many cases a single theorem is capable of being so translated in a variety of ways, there being at least one such translation for every factorization of the number $(k+1)n$ into a product of two factors $k+1$ and n such that k and n are positive integers.

Each of the four geometries may be divided into two parts. In one part we have the assumption H_0, namely:

H_0. The diagonal points of a complete quadrangle are noncollinear. In the other we have the assumption that these diagonal points are collinear. The consequences of this latter assumption are not developed in detail by Veblen and Young, but many of the theorems given as dependent on A, E, P, H_0 (so far as the given proofs go) are provable without the use of H_0 (cf. vol. I, p. 261, exercise). We have seen (§ 4) that H_0 is valid in $PG(k, p^n)$ when and only when the prime p is different from 2.

Now in volume I of the work named no assumptions are used except those which are valid for $PG(k, p^n)$. Hence every theorem in volume I may be translated, in the way indicated, into a theorem about Abelian groups. The same remarks may be made about certain parts of volume II, and in

particular about chapter III and the first part of chapter IV. It is thus apparent that our representation of the $PG(k, p^n)$ by means of Abelian groups carries at once a large part of the results of projective geometry into the domain of Abelian groups and that they there become theorems about Abelian groups. Thus by a single act of thought a significant extension is given to the theory of Abelian groups and a method is made apparent by which the theory may be further developed.

Note on a Certain Generalization of the Preceding Results. Let us now consider more generally an Abelian group A whose order is a power of a prime p and whose type is $(m_1, m_2, \cdots, m_{(k+1)n})$. Let us denote a set of independent generators of A by

$$\begin{array}{l} a_{01}, a_{02}, a_{03}, \cdots, a_{0n}, \\ a_{11}, a_{12}, a_{13}, \cdots, a_{1n}, \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ a_{k1}, a_{k2}, a_{k3}, \cdots, a_{kn}, \end{array}$$

these being chosen so that a_{ij} is of period $p^{m_{in+j}}$. Then every element of A may be represented uniquely in the form

$$\prod_{i=0}^{k} a_{i1}^{s_{i1}} a_{i2}^{s_{i2}} \cdots a_{in}^{s_{in}}$$

where the exponent s_{ij} is a number of the set $0, 1, 2, \cdots, p^{m_{in+j}} - 1$.

Consider the following subset of these elements, namely,

$$\prod_{i=0}^{k} a_{i1}^{\sigma_{i1}} a_{i2}^{\sigma_{i2}} \cdots a_{in}^{\sigma_{in}}$$

where each σ runs over the set $0, 1, 2, \cdots, p-1$, or more generally the exponent σ_{ij} runs over the set $lp^{\alpha_{ij}}$ for $l = 0, 1, 2, \cdots, p-1$, the fixed integer α_{ij} being non-negative and less than m_{in+j}. An element of this sort, for the fixed set of exponents σ_{ij},

$$\sigma_{ij} = l_{ij} p^{\alpha_{ij}},$$

the a_{ij} and the α_{ij} having been chosen once for all, may be uniquely represented by the symbol

$$\{\mu_0, \mu_1, \cdots, \mu_k\}$$

where μ_i $(i = 0, 1, 2, \cdots, k)$ denotes that mark of the Galois field $GF[p^n]$ which may be written in the form

$$\mu_i = l_{i1} + l_{i2}\omega + l_{i3}\omega^2 + \cdots + l_{in}\omega^{n-1},$$

ω being a fixed primitive mark of $GF[p^n]$.

Now let $\mu_0, \mu_1, \cdots, \mu_k$ be a fixed set of $k+1$ marks of the field $GF[p^n]$, at least one of them being different from zero; and consider the set of elements

$$\{\mu\mu_0, \mu\mu_1, \cdots, \mu\mu_k\}$$

where μ is a variable running over the p^n-1 non-zero marks of $GF[p^n]$. These elements generate a certain subgroup of A which we denote by the symbol $(\mu_0, \mu_1, \cdots, \mu_k)$. The same subgroup is denoted by the symbol $(\sigma\mu_0, \sigma\mu_1, \cdots, \sigma\mu_k)$ where σ is any non-zero mark of $GF[p^n]$. The total set of such subgroups we will call a normal set of subgroups of A.

The subgroups each of which is denoted by a symbol of the type $(\mu_0, \mu_1, \cdots, \mu_k)$ will be taken as the points of the geometry we are constructing. The point corresponding to the subgroup $(\mu_0, \mu_1, \cdots, \mu_k)$ will be denoted by the symbol $(\mu_0, \mu_1, \cdots, \mu_k)$, and $\mu_0, \mu_1, \cdots, \mu_k$ will be called the homogeneous coördinates of the point. In the geometry thus constructed the points are denoted by the same symbols as those employed in § 2 in constructing the geometry $PG(k, p^n)$ and the number system (the Galois field $GF[p^n]$) bears the same relation to the geometry in the new case as in the old. Hence the two geometries are abstractly the same. That is to say, the geometry constructed in this note is but another concrete representation of the abstract geometry $PG(k, p^n)$.

From this it follows that certain properties of the group A in the general case are identical with those for the special case when the type is $(1, 1, 1, \cdots)$, namely, those properties which may be expressed in terms of the points (and classes of points—lines, etc.) of the geometry $PG(k, p^n)$. For the sake of simplicity we shall deal with the special case when the group is of type $(1, 1, 1, \cdots)$; but the results will have the obvious extension indicated.

II. GROUPS OF ISOMORPHISMS OF ABELIAN GROUPS OF TYPE $(1, 1, 1, \cdots)$.

7. *Relation between the Groups $GLH\{k+1, p^n\}$ and I.* Let $G_{(k+1)n}$ as before be an Abelian group of order $p^{(k+1)n}$ and type $(1, 1, 1, \cdots)$. We denote it more simply by G when there is no danger of confusion. Let I denote the group of isomorphisms of G. As in the earlier part of § 2 we denote an element of this group by the symbol $\{x_0, x_1, x_2, \cdots, x_k\}$ where $x_0, x_1, \cdots, x_k$ are marks of the Galois field $GF[p^n]$.

Let us consider a linear homogeneous transformation

$$x_i' = \sum_{j=0}^{k} a_{ij}x_j, \qquad (i = 0, 1, 2, \cdots, k),$$

on the marks of this symbol, the coefficients a_{ij} being marks of $GF[p^n]$ and

the determinant $|a_{ij}|$ of this transformation being different from zero. If $\{x_0, x_1, \cdots, x_k\}$ runs over all the elements of the group G it is clear that $\{x_0', x_1', \cdots, x_k'\}$ likewise runs over all these elements. The transformation thus establishes a one-to-one correspondence of the elements of the group to its elements in some order. In each of these the identity corresponds to itself. Moreover, if $\{\mu_0, \mu_1, \cdots, \mu_k\}$ and $\{\nu_0, \nu_1, \cdots, \nu_k\}$ corresponds respectively to $\{\mu_0', \mu_1', \cdots, \mu_k'\}$ and $\{\nu_0', \nu_1', \cdots, \nu_k'\}$, then the product $\{\mu_0 + \nu_0, \cdots, \mu_k + \nu_k\}$ of the first pair of elements corresponds to the product $\{\mu_0' + \nu_0', \cdots, \mu_k' + \nu_k'\}$ of the corresponding (second) pair. Hence the correspondence of elements brought about by the given linear substitution effects an isomorphism of the group with itself. It is obvious that two distinct transformations effect different isomorphisms. Now the totality of linear homogeneous transformations of the given type constitutes the general linear homogeneous group $GLH\{k+1, p^n\}$ on $k+1$ indices with coefficients in the Galois field $GF[p^n]$. It is well known (and easily proved) that the order of this group is

$$(p^{(k+1)n} - 1)(p^{(k+1)n} - p^n)(p^{(k+1)n} - p^{2n}) \cdots (p^{(k+1)n} - p^{kn}).$$

This is a factor of the order

$$(p^{(k+1)n} - 1)(p^{(k+1)n} - p)(p^{(k+1)n} - p^2) \cdots (p^{(k+1)n} - p^{(k+1)n-1})$$

of the group I of isomorphisms of G; and it is a proper factor except when $n = 1$. Hence we have a proof of the known result that $GLH\{k+1, p^n\}$ is a subgroup of I; it is a proper subgroup when and only when $n > 1$.

Let us consider more closely isomorphisms of G with itself which are effected by the named $GLH\{k+1, p^n\}$. Let $\{\mu_0, \mu_1, \cdots, \mu_k\}$ be any element of G other than the identity and let $\{\mu_0', \mu_1', \cdots, \mu_k'\}$ be the element to which it corresponds under a given substitution belonging to $GLH\{k+1, p^n\}$. Then the element $\{\mu\mu_0, \mu\mu_1, \cdots, \mu\mu_k\}$ corresponds to the element $\{\mu\mu_0', \mu\mu_1', \cdots, \mu\mu_k'\}$ under the same substitution. Hence the subgroup $(\mu_0, \mu_1, \cdots, \mu_k)$ corresponds to the subgroup $(\mu_0', \mu_1', \cdots, \mu_k')$. Therefore every substitution in the group $GLH\{k+1, p^n\}$ effects an isomorphism of G with itself such that every subgroup of the corresponding normal set of subgroups corresponds to a subgroup of this set. Moreover, the multiplication of each coefficient a_{ij} in the transformation by one and the same non-zero mark ρ of the field gives a new transformation in which the correspondence of subgroups of the normal set as subgroups is unaltered while any other modification of the transformation, resulting in another transformation belonging to the group $GLH\{k+1, p^n\}$ leads to a different correspondence of the subgroups as such.

Now the group $GLH\{k+1, p^n\}$ has $(p^n-1, 1)$ isomorphism with the group $P(k, p^n)$ formed from the substitutions in $GLH\{k+1, p^n\}$ by treating $x_0, x_1, \cdots, x_k$ as the homogeneous coördinates in $PG(k, p^n)$, so that a substitution is now unchanged by multiplying each of its coefficients by one and the same non-zero mark ρ of the field. This group $P(k, p^n)$ is the projective group in $PG(k, p^n)$. From the result of the previous paragraph it follows that each substitution of the group $P(k, p^n)$ carries a subgroup of the normal set of subgroups into such a subgroup. Expressed geometrically this means that it transforms among themselves the points of the $PG(k, p^n)$.

When viewed geometrically, it is obvious that the group $P(k, p^n)$ also transforms planes into planes, 3-spaces into 3-spaces, and so on—facts which might be expressed also in the language of group theory. Thus a given substitution of $P(k, p^n)$ makes any given group generated by two subgroups of a normal set correspond to a group generated by two such subgroups; it also makes any given subgroups generated by three subgroups of the normal set correspond to a subgroup generated by three such subgroups; and so on.

8. *Analytical Representations of the Group I of Isomorphisms of G.* Let us consider the more general transformation

$$x_i' = \sum_{s=1}^{n} \sum_{j=0}^{k} a_{ijs} x_j^{p^{n-s}}, \qquad (i = 0, 1, 2, \cdots, k),$$

where the coefficients a_{ijs} are marks of $GF[p^n]$ such that these transformation equations have a unique solution for the symbols x_i in terms of the symbols x_i'. If

$$x_i' = \sum_{s=1}^{n} \sum_{j=0}^{k} b_{ijs} x_j^{p^{n-s}}, \qquad (i = 0, 1, 2, \cdots, k),$$

is a second transformation of the same kind, then the product of the two may be written in the form

$$\begin{aligned} x_i' &= \sum_{s=1}^{n} \sum_{j=0}^{k} a_{ijs} \Big(\sum_{\sigma=1}^{n} \sum_{\lambda=0}^{k} b_{j\lambda\sigma} x_\lambda^{p^{n-\sigma}}\Big)^{p^{n-s}} \\ &= \sum_{s=1}^{n} \sum_{j=0}^{k} a_{ijs} \Big(\sum_{\sigma=1}^{n} \sum_{\lambda=0}^{k} b_{j\lambda\sigma}^{p^{n-s}} x_\lambda^{p^{2n-s-\sigma}}\Big) \\ &= \sum_{\sigma=1}^{n} \sum_{\lambda=0}^{k} \sum_{s=1}^{n} \sum_{j=1}^{k} a_{ijs} b_{j\lambda\sigma}^{p^{n-s}} x_\lambda^{p^{2n-s-\sigma}} \\ &= \sum_{t=1}^{n} \sum_{l=0}^{k} \alpha_{ilt} x_l^{p^{n-t}}, \qquad (i = 0, 1, 2, \cdots, k), \end{aligned}$$

the α's being defined in a way which is obvious from a comparison of the last two members of the equation in the light of the fact that $x_\lambda^{p^n} = x_\lambda$. Thus

the product of two transformations of the class in consideration belongs also to the class. The named class of transformations therefore constitutes a group. This we shall call the group T. We shall prove that T, when interpreted as in the next paragraph, is identical with the group I of isomorphisms of G with itself. This result is known already for the case $k=0$ and for the case $n=1$.

Let $\{\mu_0', \mu_1', \cdots, \mu_k'\}$ and $\{\nu_0', \nu_1', \cdots, \nu_k'\}$ be the elements corresponding to $\{\mu_0, \mu_1, \cdots, \mu_k\}$ and $\{\nu_0, \nu_1, \cdots, \nu_k\}$ respectively under the given transformation with coefficients a_{ijs}. Then under the same transformation we have

$$\begin{aligned} \mu_i' + \nu_i' &= \sum_{s=1}^{n} \sum_{j=0}^{k} a_{ijs}(\mu_j^{p^{n-s}} + \nu_j^{p^{n-s}}) \\ &= \sum_{s=1}^{n} \sum_{j=0}^{k} a_{ijs}(\mu_j + \nu_j)^{p^{n-s}}, \qquad (i=0,1,2,\cdots,k). \end{aligned}$$

Hence $\{\mu_0' + \nu_0', \cdots, \mu_k' + \nu_k'\}$ corresponds to $\{\mu_0 + \nu_0, \cdots, \mu_k + \nu_k\}$ under the same transformation. Thence we see that if two given elements of G correspond respectively to two other given elements of G under a given transformation of T, then under the same transformation the product of the first pair of elements of G corresponds to the product of the second pair. Hence the substitution sets up an isomorphism of G with itself. Hence T is contained in the group I of isomorphisms of G. It remains to show that every element of I is in T.

For the latter purpose it is convenient to represent the group T in a different form.* Let ω be a primitive mark of $GF[p^n]$. Then any mark of $GF[p^n]$ may be written in the form

$$\gamma_0 + \gamma_1\omega + \gamma_2\omega^2 + \cdots + \gamma_{n-1}\omega^{n-1}$$

where each γ_i is a mark of $GF[p]$ and hence is an integer taken modulo p. Then we may write

$$x_i = \sum_{\lambda=0}^{n-1} \xi_{i\lambda}\omega^\lambda, \quad x_i' = \sum_{\lambda=0}^{n-1} \xi'_{i\lambda}\omega^\lambda, \quad a_{ijs} = \sum_{\lambda=0}^{n-1} a_{ijs\lambda}\omega^\lambda,$$

where the $\xi_{i\lambda}$, $\xi'_{i\lambda}$, $a_{ijs\lambda}$ are integers taken modulo p. Then the transformation τ of T, which has the coefficients a_{ijs}, may be written in the form

---

* The argument here is similar to that employed on pp. 69-70 of Dickson's *Linear Groups.*

$$\sum_{\lambda=0}^{n-1} \xi'_{i\lambda}\omega^{\lambda} = \sum_{s=1}^{n} \sum_{j=0}^{k} \sum_{\lambda=0}^{n-1} a_{ijs\lambda}\omega^{\lambda}\Big(\sum_{\mu=0}^{n-1} \xi_{j\mu}\omega^{\mu}\Big)^{p^{n-s}}$$
$$= \sum_{s=1}^{n} \sum_{j=0}^{k} \sum_{\lambda=0}^{n-1} \sum_{\mu=0}^{n-1} a_{ijs\lambda}\xi_{j\mu}{}^{p^{n-s}}\omega^{\mu p^{n-s}+\lambda}$$
$$= \sum_{s=1}^{n} \sum_{j=0}^{k} \sum_{\lambda=0}^{n-1} \sum_{\mu=0}^{n-1} a_{ijs\lambda}\xi_{j\mu}\omega^{\mu p^{n-s}+\lambda}, \qquad (i=0,1,2,\cdots,k).$$

Now every power of ω can be expressed linearly in terms of $\omega^0, \omega^1, \omega^2, \cdots, \omega^{n-1}$ with coefficients which are integers taken modulo p, since ω satisfies an equation of degree n with coefficients which are integers taken modulo p. On effecting this reduction we may write the last equation in the form

$$\sum_{s=1}^{n-1} \xi'_{i\sigma}\omega^{\sigma} = \sum_{\mu=0}^{n-1} \sum_{\sigma=1}^{n-1} \sum_{j=0}^{k} \alpha_{ij\mu\sigma}\xi_{j\mu}\omega^{\sigma}, \qquad (i=0,1,2,\cdots,k),$$

where the $\alpha_{ij\mu\sigma}$ are integers taken modulo p. Equating coefficients of like powers of ω we have

$$\xi'_{i\lambda} = \sum_{\mu=0}^{n-1} \sum_{j=0}^{k} \alpha_{ij\mu\lambda}\xi_{j\mu}, \quad (i=0,1,2,\cdots,k; \quad \lambda=0,1,2,\cdots,n-1).$$

Thus we have a linear transformation on the $(k+1)n$ quantities $\xi_{i\lambda}$, the coefficients of the transformation being integers taken modulo p. Since the x_i are uniquely expressible in terms of the x_i' it follows that the $\xi_{i\lambda}$ are uniquely expressible in terms of the $\xi'_{i\lambda}$ and thence that the transformation on the ξ's is non-singular.

Now the totality of such linear transformations on the $\xi_{i\lambda}$ is simply isomorphic with the group I of isomorphisms of G, as we see from the result at the end of the second paragraph of § 7 with n taken equal to 1. Hence in order to complete the proof that T is the group of isomorphisms of G it is sufficient to prove that each non-singular transformation on the $\xi_{i\lambda}$, such as the foregoing, is equivalent to a corresponding transformation in T.

In order to attain this end let the last foregoing transformation now be any non-singular linear transformation on the $\xi_{i\lambda}$ with coefficients which are integers taken modulo p. Change λ to σ, in the resulting equation (for fixed σ) multiply both sides by ω^{σ}, then sum as to σ from 0 to $n-1$. Thus we have the next preceding system of equations. From it we can go to the one which next precedes it provided that we are able to write

$$\sum_{\mu=0}^{n-1} \sum_{\sigma=1}^{n-1} \sum_{j=0}^{k} \alpha_{ij\mu\sigma}\xi_{j\mu}\omega^{\sigma} = \sum_{s=1}^{n} \sum_{j=0}^{k} \sum_{\lambda=0}^{n-1} \sum_{\mu=0}^{n-1} a_{ijs\lambda}\xi_{j\mu}\omega^{\mu p^{n-s}+\lambda}, \quad (i=0,1,2,\cdots,k),$$

where the coefficients $a_{ijs\lambda}$ are integers taken modulo p. If we have this

relation we can readily continue the reverse transformations through the equations written till we reach a transformation in the group T and having the coefficients a_{ijs}, these being marks in $GF[p^n]$. Hence, in order to show that every non-singular linear transformation on the $\xi_{i\lambda}$ (of the type in consideration) leads to a transformation of the group T it is sufficient to prove the existence of the integers $a_{ijs\lambda}$ modulo p such that the last foregoing system of equations reduces to an identity in the $\xi_{j\mu}$. For this purpose it is necessary and sufficient to show that integers $a_{ijs\lambda}$ modulo p exists such that the equation

$$\sum_{\sigma=0}^{n-1} \alpha_{ij\mu\sigma}\omega^{\sigma} = \sum_{s=1}^{n} \sum_{\lambda=0}^{n-1} a_{ijs\lambda}\omega^{\mu p^{n-s}+\lambda}$$

is valid for each set of values i, j, μ. Let us write

$$\omega^{\mu p^{n-s}+\lambda} = \sum_{\sigma=0}^{n-1} \rho_{\mu\sigma s\lambda}\omega^{\sigma},$$

where the coefficients $\rho_{\mu\sigma s\lambda}$ are integers taken modulo p. Then for the existence of the quantities $a_{ijs\lambda}$ it is necessary and sufficient that we have the relations

$$\sum_{s=1}^{n} \sum_{\lambda=0}^{n-1} \rho_{\mu\sigma s\lambda}\, a_{ijs\lambda} = \alpha_{ij\mu\sigma}$$

for every i, j, μ, σ. If i and j are held fixed, these become n^2 equations in the n^2 unknown quantities $a_{ijs\lambda}$, $s = 1, 2, \cdots, n$, $\lambda = 0, 1, \cdots, n-1$. In order that they shall have a solution it is sufficient that their determinant D shall not vanish modulo p.

In order to prove that D does not vanish modulo p we shall show that we are led to a contradiction if we suppose that $D \equiv 0 \bmod p$. If $D \equiv 0 \bmod p$ then integers $t_{s\lambda}$ exist, not all congruent to zero modulo p, such that

$$\sum_{s=1}^{n} \sum_{\lambda=0}^{n-1} t_{s\lambda}\rho_{\mu\sigma s\lambda} = 0, \qquad (\mu, \sigma = 0, 1, 2, \cdots, n-1).$$

For fixed σ multiply both members by ω^{σ}; then, summing as to σ we have a result which may be put in the form

$$\sum_{s=1}^{n} \sum_{\lambda=0}^{n-1} t_{s\lambda} \sum_{\sigma=0}^{n-1} \rho_{\mu s\lambda}\omega^{\sigma} = 0, \qquad (\mu = 0, 1, 2, \cdots, \mu - 1);$$

or, in view of the definition of the quantities ρ,

$$\sum_{s=1}^{n} \sum_{\lambda=0}^{1-n} t_{s\lambda}\omega^{\mu p^{n-s}+\lambda} = 0;$$

or,

$$\sum_{s=1}^{n} \omega^{\mu p^{n-s}} \left(\sum_{\lambda=0}^{n-1} t_{s\lambda}\omega^{\lambda} \right) = 0, \qquad (\mu = 0, 1, \cdots, n-1).$$

Now no given one of the sums in the parenthesis can be zero unless every $t_{s\lambda}$ in that sum is zero. Hence, since not every $t_{s\lambda}$ is zero, one at least of these sums in the parenthesis is different from zero. Then the consistency of the foregoing system of equations requires that the determinant

$$\Delta = \begin{vmatrix} 1 & 1 & 1 & \cdots\cdots & 1 \\ \omega^{p^{n-1}} & \omega^{p^{n-2}} & \omega^{p^{n-3}} & \cdots\cdots & \omega^{p} \\ \omega^{2p^{n-1}} & \omega^{2p^{n-2}} & \omega^{2p^{n-3}} & \cdots\cdots & \omega^{2p} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \omega^{(n-1)p^{n-1}} & \omega^{(n-1)p^{n-2}} & \omega^{(n-1)p^{n-3}} & \cdots\cdots & \omega^{(n-1)p} \end{vmatrix}$$

shall vanish. But this determinant is, apart from a constant factor, equal to a product of factors each of which is of the form

$$\omega^{p^{n-\alpha}} - \omega^{p^{n-\beta}}, \qquad (\alpha, \beta = 1, 2, \cdots, n-1, \quad \alpha \neq \beta).$$

But, since ω is a primitive mark of $GF[p^n]$, no one of these factors can vanish. Hence $\Delta \neq 0$ in $GF[p^n]$. We have been led to this contradiction by assuming that $D \equiv 0 \bmod p$. Hence this congruence is not valid.

Summing up the argument, we have the following result:

The group T, defined and interpreted at the beginning of this section, is identical with the group I of isomorphisms of the Abelian group G of type $(1, 1, 1, \cdots)$.

If the group G is of order p^m then we have a different analytical representation of the group I of isomorphisms of G for each factorization of m in the form $(k+1)n$. For the group I itself we have the simplest representation when $n = 1$. The different possible representations, however, will furnish varying information (as we shall see later) concerning various subgroups of I.

9. *On Certain Subgroups of I.* When the group I of isomorphisms of G is written in the form of the transformation group T in $GF[p^n]$, certain interesting classes of subgroups become obvious. To construct the first one of these classes we proceed as follows. Let d be any divisor of n and write $n = d\nu$. Then in the typical transformation of T put a_{ijs} equal to zero when s is not divisible by d. Then the transformation takes the special form

$$x_i' = \sum_{t=1}^{\nu} \sum_{j=0}^{k} \bar{a}_{ijt} x_j^{p^{d(\nu-t)}}, \qquad (i = 0, 1, 2, \cdots, k).$$

The product of this transformation by another of the same form may be written as a transformation of this form, the method of reduction being the same as that employed at the beginning of § 8. The named transformations therefore form a group T_d which is a subgroup of the group T, and hence (under the interpretation used in § 8) a subgroup of the group I of isomorphisms of G. It is obvious that T_1 is identical with T.

The group T_d thus formed is a generalization of the Betti-Mathieu group (see Dickson's *Linear Groups*, pp. 64-70). Just as the Betti-Mathieu group may be identified with a linear homogeneous group (Dickson, *l. c.*, p. 69), so can its generalization T_d be similarly identified with a like group. The argument is a generalization of one employed in the preceding section, whence it is sufficient merely to outline it.

Let ω be a primitive mark of $GF[p^n]$. Then any mark of $GF[p^n]$ may be written in the form

$$\gamma_0 + \gamma_1\omega + \cdot \cdot \cdot + \gamma_{\nu-1}\omega^{\nu-1}$$

where each γ_i is a mark of $GF[p^d]$. Then we may write

$$x_i = \sum_{\lambda=0}^{\nu-1} \xi_{i\lambda}\omega^\lambda, \quad x_i' = \sum_{\lambda=0}^{\nu-1} \xi'_{i\lambda}\omega^\lambda, \quad \bar{a}_{ijt} = \sum_{\lambda=0}^{\nu-1} \bar{a}_{ijt\lambda}\omega^\lambda,$$

where the $\xi_{i\lambda}$, $\xi'_{i\lambda}$, $\bar{a}_{ijt\lambda}$ are marks of $GF[p^d]$. The argument now proceeds in the same way as in the previous case and we find that

$$\xi'_{i\lambda} = \sum_{\mu=0}^{\nu-1} \sum_{j=0}^{k} \bar{\alpha}_{ij\mu\lambda}\xi_{j\mu}, \qquad (i = 0, 1, \cdot \cdot \cdot, k, \quad \lambda = 0, 1, \cdot \cdot \cdot, \nu - 1),$$

where the $\bar{\alpha}_{ij\mu\lambda}$ are marks of $GF[p^d]$. Thus a given transformation in T_d can be put into the form just written. Conversely, any transformation of the latter form can be put into the form of a transformation of T_d, the method of proof being that employed in the preceding section.

We have thus exhibited the group T_d as a homogeneous linear group in the Galois field $GF[p^d]$.

We shall now determine certain subgroups of I yielding point transformations in $PG(k, p^n)$. Let us consider the transformation

$$x_i' = \sum_{j=0}^{k} \alpha_{ijt} x_j^{p^t}, \qquad (i = 0, 1, \cdot \cdot \cdot, k),$$

belonging to the group T of § 8. On combining this transformation with the similar transformation

$$x_i' = \sum_{j=0}^{k} \beta_{ij\tau} x_j^{p^\tau}, \qquad (i = 0, 1, \cdot \cdot \cdot, k),$$

we have

$$x_i' = \sum_{j=0}^{k} \alpha_{ijt}\Big(\sum_{\lambda=0}^{k} \beta_{j\lambda\tau} x_\lambda^{p^\tau}\Big)^{p^t}$$
$$= \sum_{\lambda=0}^{k} \sum_{j=0}^{k} \alpha_{ijt} \beta^{p^t}{}_{j\lambda\tau} x_\lambda^{p^{t+\tau}}$$
$$= \sum_{\lambda=0}^{k} \gamma_{i\lambda} x_\lambda^{p^{t+\tau}}, \qquad (i = 0, 1, \cdots, k),$$

where the exponent $\tau + t$, when not less than n, is to reduced modulo n.

From this it follows that the foregoing set of transformations forms a group Γ if the coefficients α_{ijt} are marks of $GF[p^n]$ and t varies over the set $0, 1, 2, \cdots, n-1$. If d is any divisor of n and t ranges over the multiples of d in the set $0, 1, 2, \cdots, n-1$ we have a subgroup Γ_d of the group Γ. Thus we have a group Γ_d for each divisor d of n. Evidently Γ_1 is the same as Γ. We denote by Γ_0 the group in which t has the value 0 alone, this being a linear group.

Now in any particular transformation of Γ the quantities x enter homogeneously. Hence Γ has $(p^n - 1, 1)$ isomorphism with the group of point transformations which it generates in $PG(k, p^n)$. This group of point-transformations we shall denote by $C(k, p^n)$. The subgroup corresponding to the subgroup Γ_d of Γ we shall denote by $C_d(k, p^n)$. The groups $C_d(k, p^n)$ are groups of point-transformations in $PG(k, p^n)$. The group $C_0(k, p^n)$ is identical with the projective group $P(k, p^n)$ which we encountered in § 7. We shall return in § 12 to a further study of these groups.

10. *The Holomorph of G.* The set of transformations of the form

$$x_i' = x_i + a_i, \qquad (i = 0, 1, \cdots, k),$$

where the a_i are marks of $GF[p^n]$, clearly form an Abelian group $\bar{G}$ of order $p^{(k+1)n}$ and type $(1, 1, 1, \cdots)$. It is therefore simply isomorphic with the given Abelian group G and may be taken as a representation of it. The group generated by this group and group T of § 8 is therefore a representation of the holomorph of G—a fact which generalizes a well-known result (see for instance Burnside's *Theory of Groups,* 2nd ed'n, p. 245). The holomorph of G may therefore be represented by the set of non-singular transformations each of which has the form

$$x_i' = \sum_{s=1}^{n} \sum_{j=0}^{k} a_{ijs} x_j^{p^{n-s}} + a_i, \qquad (i = 0, 1, \cdots, k),$$

where the a's are marks of $GF[p^n]$. For $n = 1$ this is a well-known result. The transformation group so defined will be represented by the symbol H.

It is well-known that the group $\bar{G}$ is a self-conjugate subgroup of H. It is therefore a self-conjugate subgroup of every subgroup of H which contains $\bar{G}$. In particular $\bar{G}$ is transformed into itself by the group T_d defined in § 9. Hence the group $\{T_d, \bar{G}\}$ is a subgroup of H of the same index as that T_d in T. We thus have a ready means of constructing it. Certain of its subgroups are obvious, namely those of the form $\{T_d, \bar{G}_i\}$ where $\bar{G}_i$ is a subgroup of $\bar{G}$. An analytical representation of $\{T_d, \bar{G}\}$ is afforded by the set of non-singular transformations

$$x_i' = \sum_{t=1}^{\nu} \sum_{j=0}^{k} \bar{a}_{ijt} x_j^{p^{d(\nu-t)}} + a_i, \qquad (i = 0, 1, \cdots, k),$$

where the a's are marks of $GF[p^n]$ and d is any factor of n.

Again we can form other subgroups of H in a similar manner by taking the groups Γ_d of § 9 and combining each of them with $\bar{G}$. The forms of the analytical representations of these groups are obvious.

11. *Certain Homogeneous Groups Suggested by T and H.* At the end of § 2 we saw that the points of the Euclidean finite geometry $EG(k+1, p^n)$ may be identified with the elements of the group $G_{(k+1)n}$. Hence the group I of isomorphisms of $G_{(k+1)n}$ may be considered as a group of point transformations in $EG(k+1, p^n)$. This suggests the derivation of homogeneous groups from T and H and their subgroups and the interpretation of these in $PG(k+1, p^n)$. Accordingly we shall consider the homogeneous group whose transformations are of the form

$$x_i' = \sum_{j=0}^{k} a_{ij1} x_j^{p^{d(\nu-1)}} + \sum_{s=2}^{\nu} \sum_{j=0}^{k} a_{ijs} x_j^{p^{d(\nu-t)}} x_{k+1}^{p^{d(\nu-1)} - p^{d(\nu-t)}} + a_i x_k^{p^{d(\nu-1)}},$$
$$(i = 0, 1, 2, \cdots, k),$$
$$x'_{k+1} = x_{k+1}^{p^{d(\nu-1)}},$$

where d is any positive integral divisor of n and $n = d\nu$ (it being understood that the second summation in the first of these equations is to be omitted when $\nu = 1$). When one of the variables x_{k+1} and x'_{k+1} is 1 (or 0) the other has the same value. Therefore the given transformation transforms the points $(x_0, x_1, \cdots, x_k, 1)$ of $EG(k+1, p^n)$ according to the same permutation as that by which the corresponding transformation in $\{T_d, \bar{G}\}$ (obtained by replacing x_{k+1} by 1) transforms the elements of $G_{(k+1)n}$ when denoted by coördinates as in § 2. The fixed $PG(k, p^n)$, namely $x_{k+1} = 0$, is transformed by the foregoing substitution according to the substitution

$$x_i' = \sum_{j=0}^{k} a_{ij1} x_j^{p^{d(\nu-1)}}, \qquad (i = 0, 1, 2, \cdots, k),$$

it being assumed that the a_{ij1} are now such that this transformation is non-

singular. The total homogeneous group whose transformations are of the form of the first foregoing substitution on $x_0, x_1, \cdots, x_{k+1}$ we shall denote by $\bar{H}_d$; the subgroup in the transformations of which each a_i is zero we shall denote by $\bar{T}_d$. These are the transformation groups on the points of $PG(k+1, p^n)$ which are suggested by the non-homogeneous groups of §§ 8-10. The case when $d=1$ deserves special attention on account of its connection with the group of isomorphisms and the holomorph of $G_{(k+1)n}$.

It is obvious that there exists in $PG(k+1, p^n)$ a similar group for every k-space and corresponding Euclidean space in this $(k+1)$-space, such k-space in the new group playing the role which the k-space $x_{k+1}=0$ plays in the group as originally defined.

III. COLLINEATION GROUPS.

12. *The Collineation Group in $PG(k, p^n)$.* We shall now prove the following theorem concerning the collineation group in $PG(k, p^n)$ and certain of its subgroups.

THEOREM I. *The collineation group $C(k, p^n)$ in $PG(k, p^n)$ is represented analytically by the homogeneous transformations*

$$\text{(A)} \qquad \rho x_i' = \sum_{j=0}^{k} \beta_{ij\tau} x_j^{p^\tau}, \qquad (i=0,1,\cdots,k, \quad \tau=0,1,\cdots,n-1),$$

where the $\beta_{ij\tau}$ are marks of $GF[p^n]$ such that the determinant

$$\Delta_\tau \equiv \begin{vmatrix} \beta_{00\tau} & \beta_{01\tau} & \cdots & \beta_{0k\tau} \\ \beta_{10\tau} & \beta_{11\tau} & \cdots & \beta_{1k\tau} \\ \cdots & \cdots & \cdots & \cdots \\ \beta_{k0\tau} & \beta_{k1\tau} & \cdots & \beta_{kk\tau} \end{vmatrix}$$

is different from zero for each value of τ. Its order is n times the order of its projective subgroup $P(k, p^n)$, or $C_0(k, p^n)$, made up of those transformations of (A) in each of which $\tau=0$ and is therefore *

$$n(p^{(k+1)n}-1)p^{kn}(p^{kn}-1)p^{(k-1)n}(p^{(k-1)n}-1)\cdots p^{2n}(p^{2n}-1)p^n$$
$$=\{n/(p^n-1)\}\prod_{i=0}^{k}(p^{(k+1)n}-p^{in}).$$

The group C is generated by C_0 and the collineation

$$\rho x_i' = x_i^p, \qquad (i=0,1,\cdots,k).$$

The last element transforms C_0 into itself.

---

* Compare Dickson's *Linear Groups*, p. 87. Our group $P(k, p^n)$ is equivalent to the group of linear fractional transformations here treated by Dickson.

If d is any proper divisor of n, then we have a subgroup $C_d(k, p^n)$ of $C(k, p^n)$ (with $C_1 \equiv C$) generated by C_0 and the collineation

$$\rho x_i' = x_i^{p^d}, \qquad (i = 0, 1, \cdots, k);$$

and C_d is of index d in C. The transformations in C_d are of the form of (A) *with the restriction on τ that it shall be confined to the multiples of d belonging to the sequence* $0, 1, \cdots, n-1$.

Those transformations in C_d whose determinants Δ_τ are $(k+1)$-th *powers in $GF[p^n]$ form a subgroup $\bar{C}_d(k, p^n)$ of C_d of index μ where μ is the greatest common divisor of $k+1$* and $p^n - 1$.

The projective group $P(k, p^n)$, considered as a permutation group on the points of $PG(k, p^n)$, is triply transitive when $k = 1$ and is doubly transitive when $k > 1$. The same property of transitivity belongs to each of the previously named groups which contains $P(k, p^n)$ as a subgroup.

The group $\bar{C}_d(k, p^n)$ is doubly transitive, when considered as a permutation group on the points of $PG(k, p^n)$.

Finally, in a special case, we have another subgroup of C defined as follows. Let $k+1$ be a divisor of n, and let σ be a fixed divisor of $n/(k+1)$. Moreover, let $k+1$ be a factor of $p^\sigma - 1$. Any multiple of σ in the set $0, 1, \cdots, n-1$ can be written in just one way in the form $\{(k+1)s + \alpha\}\sigma$ where $0 \leqq \alpha \leqq k$ and s is a non-negative integer. For every such multiple of σ form the entire set of homogeneous transformations

$$\text{(B)} \qquad \rho x_i' = \sum_{j=0}^{k} \beta_{ijs\alpha} x_j^{p^{\{(k+1)s+\alpha\}\sigma}}, \qquad (i = 0, 1, \cdots, k),$$

in which each determinant $|\beta_{ijs\alpha}|$ of a transformation (s and α being fixed for a particular determinant) is equal to ω^α times a $(k+1)$-th *power in $GF[p^n]$, ω being a primitive mark of $GF[p^n]$. The totality of these transformations forms a subgroup $H_\sigma(k, p^n)$ of C which is of index $(k+1)\sigma$ in C. Moreover H_σ is contained in C_σ and is of index $k+1$ in C_σ. The group H_σ is generated by $\bar{C}_0$ and the transformations of the form*

$$\text{(C)} \qquad \rho x_i' = \omega^{t_i} x_i^{p^{\{(k+1)s+\alpha\}\sigma}}, \qquad (i = 0, 1, \cdots, k),$$

where $t_0 + t_1 + \cdots + t_k \equiv \alpha$ mod $(k+1)$. *When considered as a permutation group on the points of $PG(k, p^n)$, the group $H_\sigma(k, p^n)$ is triply transitive when $k = 1$ and is doubly transitive when $k > 1$.*

The collineation group described in the first paragraph of the theorem is the group to which we were led in § 9 in treating the subgroups of the group T of isomorphisms of G. It is an easy step to prove that the group is

a collineation group. To prove that it contains all collineations in $PG(k, p^n)$ is more difficult. But this has been effected by Veblen * through the aid of earlier work by Veblen and Bussey and by Levi. The result stated in the first paragraph of the theorem is therefore already known.

The proof of the statement in the second paragraph is omitted since it is almost immediate.

If two substitutions in C_d have their determinants equal to $(k+1)$-th powers, then their product has its determinant equal to a $(k+1)$-th power, as one may prove easily by combining these substitutions and making use of the fact that the p-th power of a determinant D whose elements are in $GF[p^n]$ is equal to the determinant $\bar{D}$ whose elements are the p-th powers of the corresponding elements of D. This proves the existence of the subgroup named in the third paragraph of the theorem. That this subgroup is of index μ in C_d is proved in general by the same method as that employed by Dickson (*l. c.*, p. 87) for the case of the group C_0.

The transitivity properties named in the fourth paragraph are immediate consequences of the fact that there exists in $P(k, p^n)$ a transformation which carries any $k+2$ points of $PG(k, p^n)$, no $k+1$ of which are on the same $(k-1)$-space, into any like set of $k+2$ points.

To show that $\bar{C}_d$ is doubly transitive we note first that the transformation (A) carries the points $(1, 0, 0, \cdots, 0)$ and $(0, 1, 0, \cdots, 0)$ into the points

$$(\beta_{00\tau}, \beta_{10\tau}, \cdots, \beta_{k0\tau}) \text{ and } (\beta_{01\tau}, \beta_{11\tau}, \cdots, \beta_{k1\tau})$$

respectively. Call these the points C and D respectively. The transformation may be chosen so that C and D are any two assigned points of $PG(k, p^n)$. Since C and D are different points there exist integers λ and μ such that the determinant $\beta_{\lambda 0\tau}\beta_{\mu 1\tau} - \beta_{\lambda 1\tau}\beta_{\mu 0\tau}$ is different from zero. Suppose now that $k > 1$. From the transformations (A) which carry the first named points into C and D respectively choose one as follows: take $\beta_{\lambda s\tau} = 0 = \beta_{\mu s\tau}$ for $s = 2, 3, \cdots, k$; choose the remaining $\beta_{ij\tau}$ for which $j > 1$ so as to give to the determinant Δ_τ any preassigned value different from zero. It is obvious that this can be done. Hence the choice of the β's and τ can be made so that the transformation (A) thus constructed belongs to the group $\bar{C}_d$. Hence the group $\bar{C}_d(k, p^n)$ is doubly transitive when $k > 1$. It is well known (cf. Dickson, *l. c.*, p. 261) and is easily proved that it is doubly transitive when $k = 1$. Hence the group $\bar{C}_d$ is doubly transitive in all cases.

It remains to prove the statements in the last paragraph of the theorem.

To show that the system of transformations named constitute a group,

---

* *Transactions of the American Mathematical Society*, Vol. 8 (1907), pp. 366-368.

consider two transformations of the named form, in one of which s and α are replaced by s_1 and α_1 and in the other of which they are replaced by s_2 and α_2. The product of these two transformations (in one order) may be written in the form

$$\rho x_i' = \sum_{j=0}^{k} \beta_{ijs_1\alpha_1} \left\{ \sum_{\mu=0}^{k} \beta_{j\mu s_2\alpha_2} x_\mu^{p^{\{(k+1)s_2+\alpha_2\}\sigma}} \right\}^{p^{\{(k+1)s_1+\alpha_1\}\sigma}}$$
$$= \sum_{\mu=0}^{k} \left\{ \sum_{j=0}^{k} \beta_{ijs_1\alpha_1} (\beta_{j\mu s_2\alpha_2})^{p^{\{(k+1)s_1+\alpha_1\}\sigma}} \right\} x_\mu^{p^{\{(k+1)(s_1+s_2)+\alpha_1+\alpha_2\}\sigma}}$$

for $i = 0, 1, \cdot \cdot \cdot, k$. It is easy to see that the determinant of this product transformation can be written as a product of determinants in the form

$$| \beta_{ijs_1\alpha_1} | \cdot | \beta_{j\mu s_2\alpha_2} |^{p^{\{(k+1)s_1+\alpha_1\}\sigma}}.$$

Now the exponent on the second determinant is congruent to 1 modulo $k+1$ since $p^\sigma - 1$ is divisible by $k+1$. Hence the determinant of the last written transformation is of the form of a $(k+1)$-th power in $GF[p^n]$ times $| \beta_{ijs_1\alpha_1} | \cdot | \beta_{j\mu s_2\alpha_2} |$. But these two determinants (by hypothesis) are equal to $(k+1)$-th powers in the $GF[p^n]$ times ω^{α_1} and ω^{α_2} respectively. Hence the determinant of the product transformation is equal to a $(k+1)$-th power in $GF[p^n]$ times $\omega^{\alpha_1+\alpha_2}$. From this and the fact that n is a multiple of $(k+1)\sigma$ it follows that the product transformation belongs to the set of transformations defined in the last paragraph of the theorem. That set therefore forms a group H_σ. It is obviously contained in C.

It is obvious that a general transformation (B) of H_σ may be multiplied by a suitable transformation (C) so as to produce a transformation belonging to $\bar{C}_0$. From this and the fact that every transformation (C) is in H_σ it follows readily that H_σ has the named generators.

It is obvious that $H_\sigma(k, p^n)$ is a subgroup of $C_\sigma(k, p^n)$. Moreover the general transformation in C_σ has its determinant restricted to be different from zero while a like transformation in H_σ has a further restriction that the value of its determinant shall be of a certain form relative to $(k+1)$-th powers so that the possible values for the determinants of transformations in C_σ of given form are $k+1$ times as many in number as the possible values for the determinants of the corresponding transformations in H_σ. From this it follows without difficulty that H_σ is of index $k+1$ in C_σ. It is therefore of index $(k+1)\sigma$ in C.

It remains to establish the transitivity properties of the group $H_\sigma(k, p^n)$. For the case $k=1$ the group was investigated by E. Mathieu.* In particular,

* *Journal de Mathématiques*, Ser. 2, Vol. 6 (1861), pp. 241-323.

he proved (p. 264) that it is triply transitive. Hence it remains to consider the case in which $k > 1$. In this case the same argument can be used as that by means of which the double transitivity of $\bar{C}_d$ was established and with the conclusion that $H_\sigma(k, p^n)$ is doubly transitive when $k > 1$.

This completes the proof of the theorem.

The transformation groups appearing in the foregoing theorem have been interpreted in it as permutation groups on the points of $PG(k, p^n)$. But these groups transform lines into lines; hence they transform the m-spaces $PG(m, p^n)$ contained in $PG(k, p^n)$ among themselves for each value m of the set $0, 1, 2, \cdots, k-1$. (Here we are taking k to be greater than 1.) Hence they may be interpreted as permutation groups on the symbols denoting the m-spaces for each particular value of m.

In particular, the $(k-1)$-spaces are transformed among themselves. The corresponding permutation group is of the same degree as that on the points of $PG(k, p^n)$, since the number of $(k-1)$-spaces in $PG(k, p^n)$ is equal to the number of points in this k-space. In view of the principle of duality it is not difficult to show that the two permutation groups arising from $C(k, p^n)$ are identical as permutation groups; for every transformation (A) on the points of $PG(k, p^n)$ can be expressed in the form of a transformation of the same general type on the coördinates which represent in a dual way the $(k-1)$-spaces $PG(k-1, p^n)$ in $PG(k, p^n)$. Moreover, the transformations (A) themselves set up a one-to-one correspondence among the elements of $C(k, p^n)$ when interpreted on the one hand as permutations on the points of $PG(k, p^n)$ and on the other hand as permutations on the $(k-1)$-spaces in $PG(k, p^n)$. Furthermore it may be seen that this correspondence is not the identical correspondence; for there are transformations leaving fixed the $(k-1)$-space $x_k = 0$ without leaving fixed any point of $PG(k, p^n)$. Detailed evidence of this fact will appear in the next section; it is involved in the fact that both the subspace $x_k = 0$ and the corresponding Euclidean space $EG(k, p^n)$ may have all its points permuted among themselves by one and the same transformation of $C(k, p^n)$.

The results of the last paragraph may be generalized to the case of l-spaces and their duals the $(k-l-1)$-spaces. Each of these sets of spaces is permuted by the transformations of $C(k, p^n)$ and the two permutation groups thus arising are identical as permutation groups. Again the simple isomorphism which is established between them is not the identical isomorphism, except in the special case of a self-dual set of spaces. This may be seen by observing that a space of the one type may be held fixed while no space of the other type is held fixed.

Hence we have the following theorem:

THEOREM II. *The collineation group $C(k, p^n)$ (when $k > 1$) transforms the $(k-1)$-spaces $PG(k-1, p^n)$ in $PG(k, p^n)$ according to the same permutation group as that according to which it transforms the points of $PG(k, p^n)$; it sets up a simple isomorphism of this permutation group with itself which is different from the identical isomorphism. More generally it sets up a like correspondence between two identical permutation groups the letters of one of which are the symbols for the l-spaces of $PG(k, p^n)$ while the letters of the other are the symbols for the dual $(k-l-1)$-spaces (except that the isomorphism may be identical in the case of self-dual spaces). These several permutation groups (of different degrees) are all simply isomorphic since each of them is simply isomorphic with $C(k, p^n)$ itself.*

It is obvious that similar results may be established for each of the subgroups of $C(k, p^n)$ described in theorem I. Of particular interest is the corresponding theorem for the case of the projective group $P(k, p^n)$. Thus *theorem II becomes a new theorem of interest if throughout it we replace $C(k, p^n)$ by $P(k, p^n)$ wherever the former occurs.*

For the case when $k > 1$ the lines of $PG(k, p^n)$ are permuted among themselves by $P(k, p^n)$, or $C(k, p^n)$, according to a transitive group, since any $k+2$ points no $k+1$ of which are on a $(k-1)$-space may be transformed into such a set of $k+2$ points by either of the named groups. If $k > 2$ the $PG(k, p^n)$ has pairs of intersecting lines and pairs of lines which do not intersect: since a pair of one of these sorts can not be transformed into a pair of the other sort, it follows that this permutation group on the lines of $PG(k, p^n)$ can not be doubly transitive when $k > 2$. When $k = 2$ the lines are transformed according to the same permutation group as the points, the latter being the dual of the former in this case. Hence the lines of $PG(2, p^n)$ are transformed among themselves according to a doubly transigroup both by $C(2, p^n)$ and by $P(2, p^n)$.

More generally it may be shown in the same way that the m-spaces $PG(m, p^n)$ in $PG(k, p^n)$, when $m < k$ and $k > 1$, are permuted according to a transitive group by either $P(k, p^n)$ or $C(k, p^n)$. If $0 < m < \frac{1}{2}k$ this group is simply transitive since there exist two sorts of pairs of m-spaces, namely, pairs in which the two spaces intersect and those in which they do not intersect, and a pair of one sort can not be transformed into a pair of the other sort by either group in consideration. Thence by means of the principle of duality it is seen that this permutation group is also simply transitive when $\frac{1}{2}k < m < k-1$. We have to consider further the case when k is even and

$m = \frac{1}{2}k$. Since this case has already been treated when $k = 2$, we shall now suppose that $k > 2$. Then for this case we have $m \geqq 2$. It is clear, then, that there exist again two sorts of pairs of m-spaces, namely, pairs in which the elements have an $(m-1)$-space in common and pairs in which the common elements constitute a space of fewer dimensions. Since a pair of one of these sorts can not be transformed into a pair of the other sort by either $P(k, p^n)$ or $C(k, p^n)$ we conclude in this case also that the permutation group on the m-spaces as symbols is simply transitive.

We shall now show that the permutation group generated in the m-spaces by $P(k, p^n)$, and hence that generated by $C(k, p^n)$, is primitive. Since the group is doubly transitive when $m = 0$ or $k - 1$ we may confine ourselves to the case in which $0 < m < k - 1$. We assume that the group is imprimitive and show that we are thus led to a contradiction. Since the m-spaces in any given $(m+1)$-space of $PG(k, p^n)$ are permuted among themselves in a doubly transitive way by the subgroup which leaves this $(m+1)$-space invariant, it follows that the m-spaces in any given $(m+1)$-space must all belong to the same set of imprimitivity. Thence it follows that the set of imprimitivity to which any given m-space M belongs must contain all the m-spaces included in the totality of $(m+1)$-spaces each of which contains M. If $m + 1 < k$ fix attention on all the $(m+1)$-spaces containing M and lying in one and the same $(m+2)$-space, and also all the $(m+1)$-spaces in this $(m+2)$-space and containing any m-space already obtained by this process of construction. Since every two $(m+1)$-spaces in the $(m+2)$-space contain an m-space in common it follows that the named process brings into consideration all the $(m+1)$-spaces in the given $(m+2)$-space. Hence every m-space in the $(m+2)$-space belongs to the same set of imprimitivity as M itself. If $m + 2 < k$ one can prove in a similar manner that the set of imprimitivity containing M contains also all m-spaces in a given $(m+3)$-space containing the given $(m+2)$-space; and so on. Hence the given set of imprimitivity contains all the m-spaces in $PG(k, p^n)$. Since this is impossible for a set of imprimitivity, we conclude that the permutation group in question is primitive.

Gathering up the results, we have the following theorem:

THEOREM III. *When $k > 1$ the collineation group $C(k, p^n)$, or its projective subgroup $P(k, p^n)$, transforms the m-spaces of $PG(k, p^n)$, $m < k$, according to a primitive permutation group; this group is doubly transitive when $m = 0$ or $k - 1$, otherwise it is simply transitive.*

From theorems II and III and from the groups $C_d(k, p^n)$ of theorem I we have the following theorem as an obvious corollary:

THEOREM IV. *There is no upper limit K to the number of primitive groups (of varying degrees) in a set of primitive groups each group of which is simply isomorphic with each of the others in the set. For every integer L there exist integers $s[t]$ such that the number of doubly transitive [triply transitive] groups of degree $s[t]$ is greater than L.*

13. *Collineation Groups Leaving Invariant an $EG(k, p^n)$.* The groups described in theorem I of § 12 obviously have corresponding subgroups each of which leaves invariant a $PG(k-1, p^n)$ in $PG(k, p^n)$. The points of $PG(k, p^n)$, not in a particular $PG(k-1, p^n)$ contained in it, form a Euclidean finite geometry of p^{kn} points; it is denoted by $EG(k, p^n)$. The named subgroups, leaving invariant a $PG(k-1, p^n)$, obviously transform among themselves the points of the corresponding $EG(k, p^n)$. Without real loss of generality we take the fixed $PG(k-1, p^n)$ to be that defined by the equation $x_0 = 0$. We then use $EG(k, p^n)$ for the corresponding Euclidean finite geometry. Concerning the named subgroups to which we are thus led, we have the following theorem which we shall now prove:

THEOREM I. *The collineation group $C(k, p^n)$ has a subgroup $EC(k, p^n)$ whose transformations may be represented analytically in the form*

$$(\bar{A}) \qquad \begin{aligned} \rho x_0' &= \beta_\tau x_0^{p^\tau}, && (\beta_\tau \neq 0), \\ \rho x_i' &= \sum_{j=0}^{k} \beta_{ij\tau} x_j^{p^\tau}, && (i = 1, 2, \cdots, k), \end{aligned}$$

where τ runs over the sequence $0, 1, 2, \cdots, n-1$. Its order is n times the order of its subgroup $EP(k, p^n)$, or $EC_0(k, p^n)$, made up of those transformations of $(\bar{A})$ in each of which $\tau = 0$ and is therefore

$$np^{kn} \prod_{i=0}^{k-1} (p^{kn} - p^{in}).$$

The group EC is generated by EC_0 and the collineation

$$\rho x_i' = x_i^p, \qquad (i = 0, 1, 2, \cdots, k).$$

The last element transforms EC_0 into itself.

If d is any proper divisor of n, then we have a subgroup $EC_d(k, p^n)$ of $EC(k, p^n)$ (with $EC_1 \equiv EC$) generated by EC_0 and the collineation

$$\rho x_i' = x_i^{p^d}, \qquad (i = 0, 1, 2, \cdots, k);$$

and EC_d is of index d in EC. The transformations in EC_d are of the form

of ($\bar{\mathrm{A}}$) *with the restriction on* τ *that it shall be confined to the multiples of* d *belonging to the sequence* $0, 1, 2, \cdots, n-1$.

Those transformations in EC_d *whose determinants are* $(k+1)$*-th powers in* $GF[p^n]$ *form a subgroup* $E\bar{C}_d(k, p^n)$ *of* EC_d *of index* μ *where* μ *is the greatest common divisor of* $k+1$ *and* p^n-1.

The group $EP(k, p^n)$, *considered as a permutation group on the* p^{kn} *points of* $EG(k, p^n)$, *is doubly transitive. Moreover, it is triply transitive when* $k > 1$ *and* $p^n = 2$. *The same property of transitivity belongs to each of the previously named groups which contains* $EP(k, p^n)$ *as a subgroup.*

Considered as a permutation group on the p^{kn} *points of* $EG(k, p^n)$, *the group* $E\bar{C}_d(k, p^n)$ *is doubly transitive when* $k > 1$ *and also when* $k = 1$ *and* $p = 2$; *it is singly transitive when* $k = 1$ *and* p *is an odd prime. This singly transitive group is primitive.*

Finally, in a special case, we have another subgroup of EC *defined as follows. Let* $k+1$ *be a divisor of* n *and let* σ *be a fixed divisor of* $n/(k+1)$. *Moreover, let* $k+1$ *be a divisor of* $p^\sigma - 1$. *Any multiple of* σ *in the set* $0, 1, 2, \cdots, n-1$ *can be written in just one way in the form* $\{(k+1)s+\alpha\}\sigma$ *where* $0 \leqq \alpha \leqq k$ *and* s *is a non-negative integer. For every such multiple of* σ *form the entire set of homogeneous transformations*

$$(\bar{\mathrm{B}}) \qquad \begin{aligned} \rho x_0' &= \beta_{s\alpha} x_0^{p^{\{(k+1)s+\alpha\}\sigma}}, && (\beta_{s\alpha} \neq 0), \\ \rho x_i' &= \sum_{j=0}^{k} \beta_{ijs\alpha} x_j^{p^{\{(k+1)s+\alpha\}\sigma}}, && (i = 1, 2, \cdots, k), \end{aligned}$$

in which each determinant of a transformation is equal to ω^α *times a* $(k+1)$*-th power in* $GF[p^n]$, ω *being a primitive mark of* $GF[p^n]$. *The totality of these transformations forms a subgroup* $EH_\sigma(k, p^n)$ *of* EC *which is of index* $(k+1)\sigma$ *in* EC. *Moreover,* EH_σ *is contained in* EC_σ *and is of index* $k+1$ *in* EC_σ. *The group* EH_σ *is generated by* $E\bar{C}_0$ *and the transformations of the form*

$$(\bar{\mathrm{C}}) \qquad \rho x_i' = \omega^{t_i} x_i^{p^{\{(k+1)s+\alpha\}\sigma}}, \qquad (i = 0, 1, 2, \cdots, k),$$

where $t_0 + t_1 + t_2 + \cdots + t_k \equiv \alpha \bmod (k+1)$. *When considered as a permutation group on the* p^{kn} *points of* $EG(k, p^n)$, *the group* $EH_\sigma(k, p^n)$ *is doubly transitive.*

That the transformations named in the first paragraph of the theorem form a group is readily verified, as is also the fact that it is generated in the way indicated. It is also easily shown that EC_0 is invariant under transformation by the last collineation defined in the paragraph. As regards this

first paragraph of the theorem it remains to show that the order given for the group is correct. For this purpose we notice that a necessary and sufficient condition on the coefficients $\beta_{ij\tau}$ is that for each τ the determinant

$$\begin{vmatrix} \beta_{11\tau} & \beta_{12\tau} & \cdot\cdot\cdot & \beta_{1k\tau} \\ \beta_{21\tau} & \beta_{22\tau} & \cdot\cdot\cdot & \beta_{2k\tau} \\ \cdot & \cdot & \cdot\cdot\cdot & \cdot \\ \beta_{k1\tau} & \beta_{k2\tau} & \cdot\cdot\cdot & \beta_{kk\tau} \end{vmatrix}$$

shall be different from zero. The number of choices of these β's and β_τ satisfying this condition for fixed τ is known (compare theorem I of § 12) to be

$$\prod_{i=0}^{k-1}(p^{kn}-p^{in}).$$

The coefficients $\beta_{i0\tau}$ may each be chosen in p^n different ways for each value of τ; and hence the set for each value of τ may be chosen in p^{kn} different ways. Taking $\tau = 0$ we see that the number of transformations in EC_0 is the number given in the theorem. From this it follows readily that EC has the order stated.

The group EC_0 has been briefly treated by Veblen and Bussey (*l. c.*, p. 255). It is obviously equivalent to the general linear (non-homogeneous) group on k variables.

After this the proofs of the statements in the second and third paragraphs of the theorem are immediate.

To establish the transitivity properties named in the fourth paragraph note first that there is in $P(k, p^n)$ a transformation that carries any $k+2$ points of $PG(k, p^n)$, no $k+1$ of which are on the same $(k-1)$-space, into any like set of $k+2$ points and that in each of two such sets two points may be taken at will in $EG(k, p^n)$ while the remaining k points may be chosen from the $(k-1)$-space $x_0 = 0$. This transformation leaves invariant this $(k-1)$-space; hence it belongs to $EP(k, p^n)$. Hence $EP(k, p^n)$, considered as a permutation group on the points of $EG(k, p^n)$, is doubly transitive.

This transitivity property may also be established analytically and thus a verification may be had of the geometric property on which the previous proof is based. Let A and B be any two points of $EG(k, p^n)$. Then there is obviously a transformation in $EP(k, p^n)$ taking A into the point $(1, 0, 0, \cdots, 0)$. Let C be the point into which this transformation takes B. To establish the named transitivity property it is then sufficient to show that C may be taken, by a transformation of $EP(k, p^n)$, into $(1, 1, 0, 0, \cdots, 0)$ while $(1, 0, 0, \cdots, 0)$ remains invariant, or, what is equivalent, that

$(1, 1, 0, 0, \cdots, 0)$ may be so taken into the point C. The transformations which are available for this are those in which each β_{i00} is zero. Then the point $(1, 1, 0, 0, \cdots, 0)$ goes into the point $(\beta_0, \beta_{110}, \beta_{210}, \cdots, \beta_{k10})$. It is obvious that the β's may be chosen so that this is the point C. Hence the named transitivity property is established analytically.

It remains to treat further the case in which $k > 1$ and $p^n = 2$. For this purpose consider those transformations of $EP(k, 2)$ which leave fixed a given point P of $EG(k, 2)$. This group is obviously simply isomorphic with the projective group in $PG(k-1, 2)$, whence it may be seen that it is doubly transitive on the points of $EG(k, 2)$ exclusive of the point P. Hence $EP(k, 2)$ is triply transitive on the points of $EG(k, 2)$.

The remaining statement in the fourth paragraph of the theorem is now obviously true. Hence the part of the theorem which is contained in that paragraph is demonstrated.

To establish the transitivity properties named in the fifth paragraph of the theorem, let us denote any two points C and D of $EG(k, p^n)$ by

$$(\beta_\tau, \beta_{10\tau}, \beta_{20\tau}, \cdots, \beta_{k0\tau}) \text{ and } (\beta_\tau, \beta_{10\tau} + \beta_{11\tau}, \beta_{20\tau} + \beta_{21\tau}, \cdots, \beta_{k0\tau} + \beta_{k1\tau}), \quad (\beta_\tau \neq 0).$$

Since C and D are distinct by hypothesis it follows that at least one $\beta_{\mu 1\tau}$ is different from zero. Let λ be a fixed quantity such that $\beta_{\lambda 1\tau} \neq 0$. Then take $\beta_{\lambda s\tau} = 0$ when $s > 1$. Taking the quantities β, as thus defined, to be the coefficients in the transformation $(\bar{A})$ which are denoted by the same symbols, we see that the points $(1, 0, 0, \cdots, 0)$ and $(1, 1, 0, 0, \cdots, 0)$ are transformed by $(\bar{A})$ into C and D respectively. Now if $k > 1$ the remaining coefficients in the transformation can be so determined that the determinant of the transformation shall have any preassigned value. Hence these coefficients may be chosen so that the transformation belongs to the group $E\bar{C}_d(k, p^n)$. From this it follows that $E\bar{C}_d(k, p^n)$ is doubly transitive when $k > 1$. It is easy to treat analytically the case when $k = 1$ and to show that $E\bar{C}_d(1, 2^n)$ is doubly transitive while $E\bar{C}_d(1, p^n)$, for $p > 2$, is only singly transitive. To prove that this singly transitive group is primitive we observe that its elements may be denoted in non-homogeneous coördinates by the transformations $t' = \alpha t + \beta$ where α is a square in $GF[p^n]$ and β is any mark of $GF[p^n]$. Then it contains the transformation $t' = \omega^2 t$ where ω is a primitive mark of $GF[p^n]$. The corresponding permutation consists of two cycles each of order $\frac{1}{2}(p^n - 1)$. All the letters in either cycle must belong to the same set of imprimitivity if the group is imprimitive, whence it follows readily that the group is primitive.

[It may be remarked in passing that the set of transformations $t' = \alpha t + \beta$ where α runs over the λ-th powers in $GF[p^n]$ and β over all the marks of $GF[p^n]$, λ being a proper factor of $p^n - 1$, form a singly transitive group of degree p^n and order $(1/\lambda)p^n(p^n - 1)$; and that this set of groups contains other primitive groups than those named in the preceding paragraph. In particular, this group is primitive when λ is a factor of $p - 1$, as may be readily shown. There are also other conditions under which it may readily be proved that the group is primitive. There are also cases in which the group is imprimitive.]

It remains to prove the statements in the last paragraph of the theorem.

The fact that the transformations $(\bar{\mathrm{B}})$ form a group may be proved in the same way as the corresponding fact was established in the case of theorem I of § 12. The proof will therefore not be given. That EH_σ has the named generators is then proved in an obvious manner. That EH_σ has the named indexes in the groups mentioned is proved in the same way as that in which the corresponding results in theorem I of § 12 were established.

The transitivity property stated in the conclusion of the theorem may be established by the method employed in establishing the transitivity properties of $E\bar{C}_d(k, p^n)$. The proof is therefore omitted. The result for $k = 1$ is given by Mathieu (*l. c.*, p. 38).

This completes the proof of the theorem.

If the coefficients β_{i00} in $(\bar{\mathrm{A}})$, $i = 1, 2, \cdots, k$, are zero, then the point $(1, 0, 0, \cdots, 0)$ is left invariant by the transformation $(\bar{\mathrm{A}})$, and conversely. Hence we have an obvious analytical representation of that subgroup of each group in theorem I which consists of all the transformations in it which leave $(1, 0, 0, \cdots, 0)$ invariant. Moreover the transitivity properties of these subgroups follow immediately from the corresponding properties of the groups as given in the theorem. The subgroup of $EC_0(k, p^n)$ which leaves $(1, 0, 0, \cdots, 0)$ fixed is obviously equivalent to the general linear homogeneous group on k variables, as Veblen and Bussey have pointed out (*l. c.*, p. 255).

It is obvious that the group $EC(k, p^n)$ is multiply isomorphic with the group $C(k-1, p^n)$ in the $PG(k-1, p^n)$ defined by the equation $x_0 = 0$. By a comparison of the orders of these two groups it is then readily shown that the isomorphism is $p^{kn}(p^n - 1)$ to 1. In a transformation $(\bar{\mathrm{A}})$ of $EC(k, p^n)$ a variation in the coefficients β_τ, $\beta_{i0\tau}$ for $i = 1, 2, \cdots, k$ and τ fixed has no effect on the permutation in the $(k-1)$-space $x_0 = 0$; and the variation of these coefficients gives $p^{kn}(p^n - 1)$ different transformations in $EC(k, p^n)$ corresponding to a given transformation in the subspace. Corresponding to the identity in $C(k-1, p^n)$ we have therefore the $p^{kn}(p^n - 1)$ transformations

$$\rho x_0' = \beta_\tau x_0, \quad \rho x_i' = \beta_{i0\tau} x_0 + x_i, \qquad (i = 1, 2, \cdots, k),$$

in $EC(k, p^n)$. It is obvious that this carries the point $(1, 0, 0, \cdots, 0)$ to any assigned point in $EG(k, p^n)$, whence this subgroup is transitive in $EG(k, p^n)$. From this it follows that every subgroup of $EP(k, p^n)$ containing all the transformations of $EP(k, p^n)$ corresponding (in the named isomorphism) to a given subgroup of $P(k-1, p^n)$ is transitive. From this it follows that for every subgroup S of $C(k-1, p^n)$ there is a corresponding subgroup T of $EC(k, p^n)$, transitive on the p^{kn} points of $EG(k, p^n)$, the latter subgroup having with the former a $p^{kn}(p^n-1)$ to 1 isomorphism. Moreover, if the former subgroup is transitive the latter is doubly transitive, a fact which may be established as follows. The largest subgroup of T which leaves fixed one point A of $EG(k, p^n)$ contains a transformation carrying any line through A into any other line through A. Hence any given point in $EG(k, p^n)$, other than A, can be carried by a transformation of T into a point B of $EG(k, p^n)$ on any other line through A, while A itself remains fixed. Then, holding this latter line fixed, as well as the point A on it, we can take a transformation* in T which leaves point-wise invariant the subspace $x_0 = 0$ and carries B to any point C in $EG(k, p^n)$ and on the line AB. Hence the subgroup of T which leaves A fixed carries any given point of $EG(k, p^n)$ other than A to any such point. Hence the largest subgroup of T which leaves A fixed is transitive on the $p^{kn}-1$ points of $EG(k, p^n)$ other than A. Hence T itself is doubly transitive on the points of $EG(k, p^n)$. When S is intransitive it is easy to show in a similar way that T is only simply transitive.

We have thus demonstrated the following theorem, except for the statements about the primitivity of the singly transitive subgroups of $EC(k, p^n)$.

THEOREM II. *The group $EC(k, p^n)$ has a $p^{kn}(p^n-1)$ to 1 isomorphism with the group $C(k-1, p^n)$ on the points of the subspace $x_0 = 0$. The subgroup T of $EC(k, p^n)$ having a $p^{kn}(p^n-1)$ to 1 isomorphism with a given subgroup S of $C(k-1, p^n)$ and corresponding to it in the isomorphism just mentioned is a transitive group, when considered as a permutation group on the p^{kn} points of $EG(k, p^n)$. Moreover, when S is transitive, the group T is doubly transitive; otherwise it is simply transitive. When S is intransitive, a necessary and sufficient condition that the simply transitive group T is primitive is that it is generated by the largest subgroup leaving the point*

* If A is taken to be the point $(1, 0, 0, \cdots, 0)$, as it may without loss of generality, the available transformation is of the form

$$x_0' = x_0, \quad x_i' = \beta x_i, \qquad \beta \neq 0, \quad i = 1, 2, \cdots, k.$$

$(1, 0, 0, \cdots, 0)$ *fixed and any (every) single transformation whatever of T that does not leave this point fixed.*

It remains to prove the statement in the last sentence. It is an immediate consequence of the general theorem * that a necessary and sufficient condition that a transitive group G is imprimitive is that the largest subgroup of G which omits one letter is contained in a larger proper subgroup of G.

Every line in the Euclidean k-space $EG(k, p^n)$ has a point in common with the projective $(k-1)$-space $x_0 = 0$ which was excluded from $PG(k, p^n)$ in forming $EG(k, p^n)$. With a line of $EG(k, p^n)$ and a point of it not on this line we may form a Euclidean plane lying in $EG(k, p^n)$; as a plane of $PG(k, p^n)$ it contains a line in the excluded $(k-1)$-space. With such a plane and an additional point of $EG(k, p^n)$ we may form a three-space which is composed of a Euclidean three-space and a plane lying in the excluded $(k-1)$-space. It is clear that this process may be continued and that one may conclude to the existence in $EG(k, p^n)$ of a Euclidean m-space $EG(m, p^n)$ for every value m of the set $1, 2, \cdots, k-1$; and in each case the remainder of the projective space $PG(m, p^n)$ which contains $EG(k, p^n)$ lies in the excluded $(k-1)$-space $x_0 = 0$.

Now any collineation group in $EG(k, p^n)$ obviously permutes among themselves the m-spaces $EG(m, p^n)$ contained in $EG(k, p^n)$. Hence each of the named groups in theorems I and II, interpreted there as a permutation group on the points of $EG(k, p^n)$, may likewise be interpreted as a permutation group on the lines of $EG(k, p^n)$, or on its planes, or on its three-spaces, or in general on its m-spaces. The several permutation groups arising in this way from one and the same transformation group are obviously simply isomorphic each to each so that they are identical as abstract groups.

Hence we have the following theorem.

THEOREM III. *Any collineation group in $EG(k, p^n)$ may be interpreted as a permutation group on the included m-spaces $EG(m, p^n)$ for each value m of the set $1, 2, \cdots, k-1$. The several permutation groups, obtained by varying the value of m, are simply isomorphic each to each.*

We shall next prove the following theorem.

THEOREM IV. *Let T and S have the same meanings as in theorem II. If S is transitive on the points of the $(k-1)$-space $x_0 = 0$, then, the group T is transitive when interpreted as a permutation group on the lines of $EG(k, p^n)$. If S is transitive on the projective l-spaces contained in the projective $(k-1)$-*

* See Miller, Blichfeldt and Dickson's *Theory of Finite Groups*, p. 39.

space $x_0 = 0$, then T is transitive on the Euclidean $(l+1)$-spaces contained in $EG(k, p^n)$; this group T is imprimitive.

The truth of the statement contained in the second sentence of the theorem is an obvious consequence of theorems II and III. To prove the statement in the last sentence we observe first that T contains a transformation carrying one point of $EG(k, p^n)$ into any other while at the same time the projective $(k-1)$-space is left pointwise invariant. Now any Euclidean $(l+1)$-space in $EG(k, p^n)$ may be defined by a projective l-space in the subspace $x_0 = 0$ and a point of $EG(k, p^n)$, it being understood that all points of $EG(k, p^n)$ collinear with the given point and the given l-space constitute the named $(l+1)$-space. Now let A and B be two Euclidean $(l+1)$-spaces so defined and let P and Q be the points in $EG(k, p^n)$ used in thus defining them. Leaving the subspace $x_0 = 0$ pointwise invariant, take P to Q by means of a transformation belonging to T. Then holding Q fixed, take the l-space of A which is in the subspace $x_0 = 0$ into the corresponding l-space of B by means of an element of T. These two transformations taken in order carry A into B. Hence T has the required property of transitivity.

It remains to be shown that the group T is imprimitive on the named $(l+1)$-spaces. For this purpose it is sufficient to observe that all the $(l+1)$-spaces of $EG(k, p^n)$ which are based, in the way indicated, on a given l-space of the subspace $x_0 = 0$ are permuted among themselves when that l-space is left invariant and that they are transformed into a like set of $(l+1)$-spaces when the given l-space is transformed into another like l-space.

14. *Collineation Groups Leaving Other Subspaces Invariant.*

We shall now prove the following theorem:

THEOREM. *The group $C^{(l)}(k, p^n)$ consisting of all transformations of the form*

$$\text{(A)} \qquad \begin{aligned} \rho x_i' &= \sum_{j=0}^{l} \beta_{ij\tau} x_j^{p^\tau}, & (i = 0, 1, 2, \cdots, l),\\ \rho x_i' &= \sum_{j=0}^{k} \beta_{ij\tau} x_j^{p^\tau}, & (i = l+1, l+2, \cdots, k), \end{aligned}$$

where $0 \leqq l < k$, τ runs over the sequence $0, 1, 2, \cdots, n-1$, and the coefficients β are marks of $GF[p^n]$, is a collineation group in $PG(k, p^n)$ which leaves invariant the subspace $PG(k-l-1, p^n)$ defined by the equations

$$x_0 = 0,\ x_1 = 0, \cdots, x_l = 0.$$

It also leaves invariant the complementary set of $p^{(k-l)n} + p^{(k-l+1)n} + \cdots + p^{kn}$ *points in* $PG(k, p^n)$. *Its order is*

$$\frac{np^{(k-l)(l+1)n}}{p^n - 1} \cdot \prod_{i=0}^{l} (p^{(l+1)n} - p^{in}) \cdot \prod_{i=0}^{k-l-1} (p^{(k-l)n} - p^{in}).$$

The group is generated by its subgroup $C_0^{(l)}(k, p^n)$ *for which* $\tau = 0$ *and the collineation*

$$\text{(B)} \qquad \rho x_i' = x_i^p, \qquad (i = 0, 1, 2, \cdots, k).$$

The last element transforms $C_0^{(l)}(k, p^n)$ *into itself.*

For each proper divisor d *of* n *the group* $C^{(l)}(k, p^n)$ *has an obvious subgroup* $C_d^{(l)}(k, p^n)$ *of index* d *generated by* $C_0^{(l)}(k, p^n)$ *and the* d*-th power of the collineation* (B). *Moreover* $C_d^{(l)}(k, p^n)$ *has an obvious subgroup* $\bar{C}_d^{(l)}(k, p^n)$ *of index* μ *consisting of those transformations of* $C_d^{(l)}(k, p^n)$ *whose determinants are* $(k+1)$*-th powers,* μ *being the greatest common divisor of* $k+1$ *and* $p^n - 1$.

The common subgroup of $C^{(l)}(k, p^n)$ *and the group* $H_\sigma(k, p^n)$ *of theorem I of* § 12 *consists of the entire set of transformations of the form*

$$\rho x_i' = \sum_{j=0}^{l} \beta_{ijsa} x_j^{p^{\{(k+1)s+a\}\sigma}}, \qquad (i = 0, 1, 2, \cdots, l),$$

$$\rho x_i' = \sum_{j=0}^{k} \beta_{ijsa} x_j^{p^{\{(k+1)s+a\}\sigma}}, \qquad (i = l+1, l+2, \cdots, k),$$

the notation being that of theorem I of § 12 *and the determinant* $|\beta_{ijsa}|$ *being restricted as in that theorem.*

The group $C_0^{(l)}(k, p^n)$ *is transitive when interpreted as a permutation group on the set of* $p^{(k-l)n} + \cdots + p^{kn}$ *points mentioned in the first paragraph of the theorem.*

That the given set of transformations form a group leaving invariant the named subspace, and hence the complementary set of points, is obvious. To determine the order of the group we notice first that the determinant of the coefficients $\beta_{ij\tau}$ for i and j running over the set $0, 1, 2, \cdots, l$ must be different from zero; whence it follows (from a comparison with theorem I of § 12) that these coefficients can be chosen in

$$\prod_{i=0}^{l} (p^{(l+1)n} - p^{in})$$

different ways, τ remaining fixed. The coefficients $\beta_{ij\tau}$ for i and j running over the set $l+1, l+2, \cdots, k$ and τ remaining fixed can then be chosen

independently in any way so that their determinant shall be different from zero; and hence they can be chosen in

$$\prod_{i=0}^{k-l-1} (p^{(k-l)n} - p^{in})$$

different ways. Then for τ still fixed each of the remaining $(k-l)(l+1)$ coefficients β can be chosen independently in p^n ways, so that altogether this set of coefficients can be chosen in

$$p^{(k-l)(l+1)n}$$

different ways. Finally there are n values for τ. Hence the order of the group is the product of n and the three numbers just determined, all divided by $p^n - 1$, this divisor being introduced to allow for the factor of proportionality. From this it follows that the order of the group is that stated in the theorem.

That the group is generated in the way indicated is obvious.

The propositions in the second paragraph of the theorem are obvious in view of the corresponding parts of theorem I of § 13.

The proposition in the third paragraph of the theorem has an obvious demonstration in view of the proof of the corresponding part of theorem I of § 12.

Since any $k+2$ points no $k+1$ of which are on a $(k-1)$-space can be carried by the projective group into any other such set, it is obvious that an l-space may be held fiexd while any point not on it is transformed into any other such point. Thence follows readily the truth of the last proposition in the theorem.

Grundlagen der kombinatorischen Logik.

TEIL II.*

von H. B. Curry.

C. Darstellung der Kombinationen durch Kombinatoren in der Normalform.

In diesem Abschnitte gebrauchen wir gewisse Zeichen, die wir Variablen nennen wollen. Diese Variablen sind nur ein Hilfsmittel, womit wir zeigen können, dass eine gewisse Art von Vollständigkeit und Verträglichkeit des Grundgerüstes vorliegt. Sie sind nicht als Ableitungen des Grundgerüstes anzusehen. Die Ausführungen dieses Abschnitts haben daher mit der formalen Darstellung nichts zu tun, sondern sie betreffen die Verwandtschaft zwischen dieser und der gewöhnlichen Logik. Diese Variablen sind als Etwase ohne besondere Eigenschaften zu behandeln.

Die Hauptergebnisse dieses Abschnitts sind die letzten Sätze von § 1 und § 5. Unter den ersten kommt der Hauptsatz I von Abschnitt A vor; dagegen macht § 5, Satz 2 den Kern des Hauptsatzes II aus.

§ 1. *Allgemeines über Reduktion und Entsprechen; ihre Eindeutigkeit.*

Festsetzung 1. In dem Folgenden betrachten wir Ausdrücke, die aus gewissen Variablen $x_1, x_2, x_3, \cdots$ und Etwase formal aufgebaut werden, d. h. so dass die Variablen als Etwase ohne besondere Eigenschaften behandelt werden. Auf solche Ausdrücke werden die vorhergehenden Festsetzungen und Definitionen ausgedehnt.

Festsetzung 2. Wir betrachten nun ein X, das eine Kombination von Kombinatoren und Variablen ist. Wir nehmen an, dass X mit den nach I C, Def. 1, erlaubten Auslassungen von Klammern geschrieben ist, und dass alle die anderen in den vorigen Abschnitten definierten Bezeichnungen durch ihre Definitionen ersetzt sind. Dann ist X von der Form $(X_0 X_1 X_2 \cdots X_n)$, wo X_0 entweder B, C, W, K oder eine Variable x_k ist, und die X_i für $i > 0$ Ausdrücke von derselben Form wie X sind.

Inbezug auf einen solchen X setzen wir zwei Arten von Reduktionsprozessen fest, wie folgt:

1.) Wenn X_0, $B, C, W,$ oder K ist und n gross genug ist, so dürfen wir für

* Teil I erschien in diesem Journal, Bd. 52 (1930), S. 509-536.

$X_0X_1X_2$ bzw. $X_0X_1X_2X_3$ sein Äquivalent nach der betreffenden Regel B, C, W oder K ersetzen, z. B., wenn $X_0 \equiv B$,* so haben wir $X_1(X_2X_3)X_4 \cdots X_n$ anstatt $X_0X_1X_2X_3X_4 \cdots X_n$. Eine solche Ersetzung soll ein *Reduktionsprozess erster Art* heissen.

2.) Es mag sein, dass ein Bestandteil von X (d. h. ein eingeklammerter in X erscheinender Ausdruck) durch einen Reduktionsprozess erster Art umgeformt werden kann. Eine solche Umformung soll ein *Reduktionsprozess zweiter Art* heissen, wenn ein Reduktionsprozess erster Art sowohl für den Gesamtausdruck wie auch für jeden Teilausdruck, der den betrefienden einschliesst oder links von ihm steht, unmöglich ist.

Festsetzung 3. Ein Ausdruck X *reduziert* sich auf einen anderen Y, wenn durch Anwendung dieser Prozesse X in Y umgeformt wird, und zwar *im ersten Sinne,* wenn nur Prozesse erster Art nötig sind, und im *zweiten Sinne,* wenn auch Prozesse zweiter Art nötig sind. Dass X sich auf Y reduziert wird auch durch das Zeichen $X \doteq Y$ ausgedruckt.

Festsetzung 4. Es sei ein Ausdruck X_m gegeben, der x_m, aber keine x_n, $n > m$, enthält, und der ferner nicht von der Form $X_{m-1}x_m$ ist. Dann denken wir an die unendliche Zeichenfolge, welche entsteht, wenn man rechts von X_m die Variablen $x_{m+1}, x_{m+2}, \cdots$ ad infin. setzt; diese heisst die durch X_m bestimmte *Folge.* Der Teil dieser Folge, welcher einem bestimmten x_n, $n > m$, vorangeht, ist ein Ausdruck, der ein *Abschnitt* der Folge heisst. Also ist X_m selbst ein Abschnitt der durch ihn bestimmten Folge.

Festsetzung 5. Ein Ausdruck X enthält eine Variable x_m *wesentlich,* wenn für $n >$ den Index irgendeiner in X erscheinenden Variablen, und für $p \geqq 0$, in der Reduktion von $(Xx_nx_{n+1} \cdots x_{n+p})$ die Variable x_m nie ausfällt.† Z. B. der Ausdruck $B(Kx_1)x_2$ enthält x_1, wesentlich, aber nicht x_2. Die höchste wesentlich erscheinende Variable in X heisst der *Grad* von X. (Wenn keine Variable wesentlich erscheint, so heisst der Grad 0).

Festsetzung 6. In der Reduktion eines Ausdrucks X auf einen anderen Y heisst eine Variable x_n *nicht gestört,* wenn 1) X von der Form $X'x_nx_{n+1} \cdots x_{n+p}$ ist, wo X' die Variablen $x_n, x_{n+1}, \cdots, x_{n+p}$ nicht wesentlich enthält, 2) Y von einer ähnlichen Form $Y'x_nx_{n+1} \cdots x_{n+p}$ ist, 3) X' sich auf Y' reduzieren lässt. Sonst heisst eine in X erscheinende Variable gestört.

Festsetzung 7. Ein Ausdruck X *entspricht* einer Folge $\mathfrak{X}$, wenn die fol-

* Der leser soll bemerken, dass Ausdrücke wie $X \equiv Y$ und $\vdash X$. Sätze bedeuten. (s. IC).

† Natürlich soll x_m nicht in X selbst fehlen.

gende Bedingung erfüllt ist: es gibt ein $p \geqq 0$, so dass der Ausdruck $(Xx_{n+1}x_{n+2}\cdots x_{n+p})$, wo n der Grad von X ist, sich auf einen Abschnitt von $\mathfrak{X}$ reduziert, und zwar so, dass x_{n+p}, wenn $n+p>0$ ist, nicht ausgelassen wird. Wenn x_{n+q} die höchste in dieser Reduktion gestörte Variable ist (bzw. $q=0$, wenn keine nicht in X wesentlich erscheinende Variable gestört wird), so sagen wir, dass X der Folge *mit der Ordnung q* entspricht.* Endlich sprechen wir von einem Entsprechen *im ersten bzw. zweiten Sinne,* wenn die Reduktion sich im ersten bzw. zweiten Sinne vollzieht.

Festsetzung 8. Zwei Ausdrücke X und Y heissen *äquivalent* im

1) *ersten Sinne,* wenn sie denselben Grad haben und derselben Folge von lauter Variablen entsprechen,

2) *zweiten Sinne,* wenn sie denselben Grad haben, und derselben Folge von lauter Variablen mit derselben Ordnung entsprechen,

3) *dritten Sinne,* wenn sie denselben Grad haben, und derselben Folge von lauter Variablen in demselben Sinne entsprechen.

4) *vierten Sinne,* wenn sie denselben Grad haben, und derselben Folge von lauter Variablen in demselben Sinne und mit derselben Ordnung entsprechen.

Bemerkung: Die folgenden Sätze haben als Zweck den Beweis, dass, wenn eine Formel der Form $\vdash X = Y$ aus unserem Grundgerüst ableitbar ist, ein gewisser Sinn von Äquivalenz zwischen X und Y besteht. Eine gewisse Art von Übereinstimmung mit Logik und Unabhängigkeit wird dabei für die kombinatorischen Axiome gewährleistet. Dies ist die einzige solche Untersuchung dieser Abhandlung; eine allgemeine Vollständigkeits-, Widerspruchslosigkeits- oder Unabhängigkeitsuntersuchung wird von dieser Abhandlung ausgeschlossen.

Hilfssätze. Das Reduzieren ist seiner Definition nach ein eindeutiger Prozess, also haben wir leicht die folgenden Hilfsätze.

1. Wenn ein Ausdruck sich auf zwei verschiedene Ausdrücke reduziert, so reduziert einer dieser beiden sich auf den anderen.

---

* Man darf hier annehmen dass entweder $p = q$ oder $p = q + 1$ ist. Denn nach den Voraussetzungen reduziert $Xx_{n+1}x_{n+2}\cdots x_{n+p}$ sich auf ein $\mathfrak{C}x_{n+q+1}x_{n+q+2}\cdots x_{n+p}$ und zwar so, dass $x_{n+q+1}, x_{n+q+2}, \cdots, x_{n+p}$ dabei ungestört werden. Daher reduziert $Xx_{n+1}x_{n+2}\cdots x_{n+q+1}$ sich auf $\mathfrak{C}x_{n+q+1}$; und dieser ist ein Abschnitt der Folge $\mathfrak{F}$, näm., $\mathfrak{C}x_{n+q+1}x_{n+q+2}\cdots$. Wir wissen ja auch, dass $Xx_{n+1}\cdots x_{n+q}$ sich auf $\mathfrak{C}$ reduziert, aber davon können wir nicht schliessen, dass immer $p = q$ sein kann, weil $\mathfrak{C}$ nicht ein Abschnitt der Folge $\mathfrak{F}$ ist, falls x_{n+q} in der Reduction ausfällt.

2. Ein Ausdruck kann nie auf zwei verschiedene Kombinationen von Variablen reduziert werden.

3. Ein Ausdruck kann nie zwei verschiedenen Folgen von Variablen entsprechen.

4. Wenn X und Y denselben Grad n haben, und wenn ferner die zwei Ausdrücke $(Xx_{n+1}x_{n+2}\cdots x_{n+p})$ und $(Yx_{n+1}x_{n+2}\cdots x_{n+p})$ sich auf dieselbe Kombination lauter Variablen reduzieren, so sind X und Y äquivalent in dem ersten Sinne.

5. Wenn X und Y denselben Grad n haben, und wenn ferner für jedes p, wofür einer der beiden Ausdrucke $(Xx_{n+1}x_{n+2}\cdots x_{n+p})$ und $(Yx_{n+1}x_{n+2}\cdots x_{n+p})$ auf eine Kombination von lauter Variablen reduziert wird, die beiden sich auf dieselbe Kombination reduzieren, so sind X und Y in dem zweiten Sinne äquivalent.

SATZ 1. *Sind $\mathfrak{A}_1, \mathfrak{A}_2, \cdots, \mathfrak{A}_N, \mathfrak{B}_1, \mathfrak{B}_2, \cdots, \mathfrak{B}_N, X, Y$, Kombinationen von Kombinatoren und Variablen $x_1, x_2, \cdots, x_m$ derart, dass*

1) *für jedes i $(i = 1, 2, \cdots N)$ die beiden Ausdrücke $\mathfrak{A}_i$ und $\mathfrak{B}_i$ denselben Grad haben, und weiter derselben Folge mit derselben Ordnung und im ersten Sinne entsprechen,*

2) *X einer Folge von lauter Variablen entspricht,*

3) *aus den Voraussetzungen*

$$\vdash \mathfrak{A}_i = \mathfrak{B}_i \tag{1}$$

mit Benutzung nur der Eigenschaften der Gleichheit (I D) folgt, dass

$$\vdash X = Y; \tag{2}$$

dann sind X und Y äquivalent im zweiten Sinne, und zwar, wenn jedes $\mathfrak{A}_i$ und jedes $\mathfrak{B}_i$ wirklich Kombinatoren enthält, im vierten Sinne.

Beweis: Es genügt, den Satz für den Fall zu beweisen, dass X aus Y durch eine einzige Ersetzung entsteht, nämlich der Ersetzung eines in X erscheinenden $\mathfrak{A}$ durch seinen Gegenwert $\mathfrak{B}$, oder umgekehrt. Das allgemeinste Y ergibt sich aus X durch eine Reihe von solchen Ersetzungen.

Nach Hp. 2 gibt es ein n, wofür $(Xx_{m+1}x_{m+2}\cdots x_{m+n})$ sich auf eine Kombination von $x_1, x_2, \cdots, x_{m+n}$ reduziert. Ich möchte diese Kombination Z nennen, und die Ausdrücke $(Xx_{m+1}x_{m+2}\cdots x_{m+n})$ bzw. $(Yx_{m+1}x_{m+2}\cdots x_{m+n})$ mit X' und Y' abkürzen. Ich zeige zunächst, dass Y' auf Z reduziert wird, und zwar, wenn die $\mathfrak{A}_i$ und $\mathfrak{B}_i$ alle wirklich Kombinatoren enthalten, in demselben Sinne.

$\mathfrak{A}$ sei der ersetzte Ausdruck in X und $\mathfrak{B}$ sein Gegenwert, so lässt Y' sich von X' nur dadurch unterscheiden, dass in Y' $\mathfrak{B}$ die Stelle von $\mathfrak{A}$ einnimmt. Dann können im Laufe der Reduktion die folgenden drei Möglichkeiten geschehen:

I. Wir kommen zu einer Form an, worin $\mathfrak{A}$ am Anfang steht, d. h. zu einer Form

$$(\mathfrak{A}X_1X_2\cdots X_p), \tag{3}$$

wo die $X_1, X_2, \cdots, X_p$ Kombinationen von Kombinatoren und Variablen sind.

II. Ein eingeklammerter Teilausdruck, der $\mathfrak{A}$ enthält (bzw. $\mathfrak{A}$ selbst), wird als ein Ganzes durch K ausgestrichen.

III. $\mathfrak{A}$ bleibt innerhalb des Gesamtausdrucks (d. h. nicht am Anfang), bis in der Reduktion durch Prozesse zweiter Art die Reihe an es kommt, und dann steht es am Anfang eines Teilausdrucks der Form (3), wo $p \geqq 0$ ist.

Diese drei Möglichkeiten sind erschöpfend, weil Reduktion so definiert ist, dass $\mathfrak{A}$ sonst ein untrennbares Ganzes ist. Ich behandle die drei Fälle jetzt besonders.

Fall I. Nach der Voraussetzung dieses Falles reduziert X' sich auf einen Ausdruck X'' der Form (3). Dann reduziert sich Y' durch genau dieselbe Reihe von Reduktionsprozessen auf ein Y'' der Form

$$(\mathfrak{B}X_1X_2\cdots X_p), \tag{4}$$

wo die $X_1, X_2, \cdots, X_p$ dieselben Ausdrücke wie in X'' sind.

Es werde nun angenommen, der Ausdruck $\mathfrak{A}' \equiv (\mathfrak{A}x_{m+1}x_{m+2}\cdots x_{m+p})$ reduziert sich im ersten Sinne auf einen Ausdruck $\mathfrak{C}$. Dann, wenn wir überall in dieser Reduktion $x_{m+1}, x_{m+2}, \cdots x_{m+p}$ durch $X_1, X_2, \cdots X_p$ ersetzen, liefert die so entstehende Folge von Ausdrücken wieder eine Reduktion im ersten Sinne. Daher reduziert sich X'' durch Prozesse erster Art auf ein X''', welches entsteht, wenn man in $\mathfrak{C}$ die betreffenden Einsetzungen macht. Eine ähnliche Bemerkung bezieht sich auf Y''.

Nach Hp. 1 entsprechen $\mathfrak{A}$ und $\mathfrak{B}$ beide derselben Folge $\mathfrak{F}$. r sei die Ordnung, womit $\mathfrak{A}$ dem $\mathfrak{F}$ entspricht. Dann zeige ich, dass $r \leqq p$ ist. In der Tat nehmen wir das Gegenteil an. Dann reduziert der Ausdruck $(\mathfrak{A}'x_{m+p+1}\ x_{m+p+2}\cdots x_{m+r+1})$ sich auf einen Abschnitt von $\mathfrak{F}$ und zwar so, dass x_{m+p+1} gestört wird.* Unter der durch diese Reduktion erzeugten Reihe von Ausdrücken gibt es ein $(\mathfrak{C}x_{m+p+1}x_{m+p+2}\cdots x_{m+r+1})$ derart, dass die Reduktion sich bis auf diesen Ausdruck ohne Störung von $x_{m+p+1}\cdots x_{m+r+1}$ erstrickt,

* S. Festsetzung 6, Anmerkung.

während im nächsten Schritte der Reduktion x_{m+p+1} gestört wird. Also muss $\mathfrak{C}$ von der Form

$$(5) \qquad (X_0''X_1''X_2'' \cdot \cdot \cdot X_q'')$$

sein, wo entweder 1) X_0'' B oder C ist und $q < 3$ ist, oder 2) X_0'' W oder K ist und $q < 2$ ist. Nach Festsetzung 6 reduziert $\mathfrak{A}'$ sich auf dieses $\mathfrak{C}$. Dann reduziert sich X'' nach dem vorigen Absatz auf ein X''' derselben Form (5). Aber in der weiteren Reduktion eines solchen X''' könnte der Kombinator X_0'' nie verschwinden, was der Voraussetzung, dass X' sich auf Z reduziert, widerspricht.

Also gilt $r \leqq p$. Dann reduzieren sich die beiden Ausdrücke $(\mathfrak{A}'x_{m+p+1})$ und $(\mathfrak{B}'x_{m+p+1})$ auf einen Abschnitt $(\mathfrak{C}x_{m+p+1})$ von $\mathfrak{F}$, und zwar so, dass x_{m+p+1} ungestört wird. Daher reduzieren sich $\mathfrak{A}'$ und $\mathfrak{B}'$ beide auf dasselbe $\mathfrak{C}$ (Festsetzung 6). Diese Reduktion geschieht weiterhin im ersten Sinne. Nach dem vorletzten Absatz reduzieren sich dann X'' und Y'' auf ein gemeinsames X''', und zwar im ersten Sinne. Weil X' auf Z reduziert wird, so reduziert sich X''', und also Y' auf Z. Weil die einzigen Reduktionsprozesse, die in den Reduktionen von X' und Y' verschieden sind, zu der ersten Art gehören, so reduzieren X' und Y' sich auf Z in demselben Sinne.

Fall II. Durch eine Reihe von Reduktionsprozessen reduziert X' sich auf einen Ausdruck X'', der einen Teilausdruck der Form $(KX_1X_2 \cdot \cdot \cdot X_p)$ enthält, wo $\mathfrak{A}$ in X_2 enthalten ist, und zwar so, dass beim nächsten Schritte die Reduktion auf einen X''' führt, der sich vom X'' nur dadurch unterscheidet, dass der obige Teilausdruck durch $(X\,X_3 \cdot \cdot \cdot X_p)$ ersetzt ist. Genau dieselbe Reihe von Prozessen reduziert Y' auf einen Ausdruck Y'', der sich von X'' nur darin unterscheidet, dass $\mathfrak{A}$ die Stelle von $\mathfrak{B}$ einnimmt. Beim nächsten Schritte, der derselbe Prozess wie im vorigen Falle ist, kommen wir wieder auf X'''. Daher reduzieren X' und Y' sich durch dieselbe Reihe von reduktionsprozessen auf denselben Ausdruck. Infolgedessen reduzieren sie sich endlich auf dieselben Kombination, und zwar, weil die beiden Reihen von Prozessen dieselben sind, in demselben Sinne.

Fall III. Nach der Voraussetzung reduziert X' sich auf einen Ausdruck X'', der einen Teilausdruck der Form (3) enthält, und zwar so, dass die weitere Reduktion von X'' durch die Reduktion dieses Teilausdrucks fortgesetzt wird. Dann reduziert Y' sich auf ein Y'', welches sich von X'' nur darin unterscheidet, dass der Ausdruck (4) anstatt (3) erscheint.

Weil die Bedingungen von Fall I für diese Teilausdrücke (3) und (4) erfüllt sind, so reduzieren diese Teilausdrücke sich auf dieselben Kombinationen. Weil X'' und Y'' sonst identisch sind, so reduzieren X'' und Y'',

und daher auch X' und Y' sich auf denselben Ausdruck. Infolgedessen werden X' und Y' auf dieselbe Kombination von Variablen reduziert.

Wenn $\mathfrak{A}$ und $\mathfrak{B}$ wirklich Kombinatoren enthalten, so sind Reduktionsprozesse zweiter Art in den beiden Fällen erforderlich. Deshalb werden sie auf diese Kombination in demselben Sinne reduziert.

Es ist nun bewiesen, dass X' und Y' sich auf dasselbe Z reduzieren. Daraus folgt zunächst, dass X und Y denselben Grad haben; denn jede Variable, die in der Reduktion von X' verschwindet, verschwindet auch in der Reduktion von Y', und umgekehrt. Dieser Grad sei dann μ. Setzen wir in den obigen Beweis $x_{\mu+j}$ statt x_{m+j} ein, so folgt, dass die neuen X' und Y' auch auf eine gemeinsame Kombination lauter Variablen reduziert werden, wenn nur eines von den beiden sich auf eine solche Kombination reduziert. Also entsprechen X und Y derselben Folge mit derselben Ordnung (Hilfsatz 5), und auch, wenn die $\mathfrak{A}$ und $\mathfrak{B}$ wirklich Kombinatoren enthalten, in demselben Sinne, w. z. b. w.

SATZ 2. *Sind* $\mathfrak{A}_1, \mathfrak{A}_2, \cdots, \mathfrak{A}_N, \mathfrak{B}_1, \mathfrak{B}_2, \cdots, \mathfrak{B}_N, X, Y,$ *Kombinationen von Kombinatoren und Variablen* $x_0, x_1, \cdots, x_m$ *derart, dass die Bedingungen von Satz* 1 *erfüllt sind, ausser dass in Hp.* 3 *bei der Ableitung von* (2) *aus* (1) *auch Benutzung von den Regeln* B, C, W *und* K *erlaubt wird; dann sind* X *und* Y *im zweiten Sinne äquivalent.*

Beweis: Wir können von X zu Y durch eine Reihe von Schritten übergehen, wovon jeder daraus besteht, dass wir entweder eine einzige Ersetzung aus den Formeln (1) machen, oder auch eine Regel B, C, W oder K einmal anwenden. Weiter dürfen wir unter einer solchen Anwendung den folgenden Prozess verstehen: zunächst setzen wir in einer Regel (B, C, W oder K) für die X, Y (und Z, wenn es erscheint) besondere Ausdrücke ein, sodass eine Formel $\mathfrak{A} = \mathfrak{B}$ entsteht, und dann machen wir in einem schon aus X abgeleiteten Ausdruck eine Ersetzung von $\mathfrak{A}$ durch $\mathfrak{B}$ oder umgekehrt.

Jetzt betrachten wir alle die Formeln, die in diese Weise aus allen den im Uebergang von X zu Y benutzten Anwendungen der betreffenden Regeln entstehen. Fügen wir diese Formeln zu den Formeln (1) hinzu. Dann sind alle die Bedingungen von Satz 1 für die erweiterte $\mathfrak{A}_1, \mathfrak{A}_2, \cdots \mathfrak{A}_N, \mathfrak{A}_{n+1}, \cdots, \mathfrak{A}_M, \mathfrak{B}_1, \mathfrak{B}_2, \cdots, \mathfrak{B}_N, \mathfrak{B}_{N+1}, \cdots, \mathfrak{B}_M, X, Y$ erfüllt. Also folgt der Satz aus Satz 1.

Es soll bemerkt werden, dass die Nebenbedingung für den strengen Satz 1, nämlich dass alle die $\mathfrak{A}_i$ und $\mathfrak{B}_i$ wirklich Kombinatoren enthalten, für die neue $\mathfrak{A}_{N+j}$ oder $\mathfrak{B}_{N+j}$ versagen mag, sogar wenn sie für die ursprüngliche $\mathfrak{A}_i$ und $\mathfrak{B}_i$ erfüllt ist.

SATZ 3. *Wenn $\mathfrak{A}_1, \mathfrak{A}_2, \cdots, \mathfrak{A}_N, \mathfrak{B}_1, \mathfrak{B}_2, \cdots, \mathfrak{B}_N$, X, Y Kombinatoren sind, die die Hypothesen von Satz 2 erfüllen; dann sind X und Y äquivalent im vierten Sinne.*

Beweis: Die $\mathfrak{A}_i$ und $\mathfrak{B}_i$, die sowohl in den ursprünglichen Formeln (1), als auch in denen, die dazu durch die Prozesse des Beweises von Satz 2 hinzugefügt werden, erscheinen, sind Kombinatoren und enthalten daher Kombinatoren. Also folgt der Satz aus Satz 1.

SATZ 4. *Sind X, Y Kombinatoren, wofür*

1) *es folgt aus den transmutativen Axiomen mit Benutzung der Regeln B, C, W, K und den Eigenschaften der Gleichheit, dass $\vdash X = Y$,*

2) *mindestens einer der beiden einer Folge von lauter Variablen entspricht;*

dann sind X und Y äquivalent im vierten Sinne.

Beweis: folgt aus Satz 3, weil die betreffenden Axiome die Bedingungen der Formeln (1) erfüllen.

SATZ 5. *Wenn $\mathfrak{A}_1, \mathfrak{A}_2, \cdots, \mathfrak{A}_N, \mathfrak{B}_1, \mathfrak{B}_2, \cdots, \mathfrak{B}_N$, X, Y Kombinationen von Variablen und Kombinatoren sind, derart, dass*

1) *für jedes i $(i = 1, 2, \cdots, N)$ $\mathfrak{A}_i$ und $\mathfrak{B}_i$ denselben Grad haben und weiter derselben Folge mit derselben Ordnung entsprechen,*

2) *X einer Folge von lauter Variablen entspricht,*

3) *aus den Formeln*

$$(1) \qquad \vdash \mathfrak{A}_i = \mathfrak{B}_i \qquad (1 = 1, 2, \cdots, N)$$

mit Benutzung der Regeln B, C, W, K und der Eigenschaften der Gleichheit folgt, dass

$$(2) \qquad \vdash X = Y;$$

dann sind X und Y äquivalent im zweiten Sinne.

Beweis: Zunächst sehen wir sofort, dass der Satz, wenn er für den Fall bewiesen ist, dass im Hp. 3) die Benutzung nur von den Eigenschaften der Gleichheit erlaubt ist, im allgemeinen durch das Verfahren, das ich in dem Beweis von Satz 2 benutzt habe, bewiesen werden kann. Es genügt daher, den Satz für jenen Fall zu beweisen.

Der Beweis verläuft nun wie der von Satz 1. Wir setzen ohne Beschränkung der Allgemeinheit voraus, dass Y sich aus X durch eine einzige Einsetzung, die von $\mathfrak{A}$ statt $\mathfrak{B}$, ergibt. Wir definieren X', Y' und Z wie dort, und schliessen, wie folgt, dass Y' auf Z reduziert wird. Wir unterscheiden dieselben drei Fälle wie im Satz 1.

Fall I. X' und Y' reduzieren sich auf X'' bzw. Y'' von der Form (3) bzw. (4).

Es sei nun angenommen, der Ausdruck $\mathfrak{A}'$ (definiert wie im Satz 1) reduziert sich auf einen Ausdruck $\mathfrak{C}$; dann, wenn wir überall in dieser Reduktion $x_{m+1}, x_{m+2}, \cdots, x_{m+p}$ durch $X_1, X_2, \cdots, X_p$ ersetzen, so schaffen wir eine Reihe von Ausdrücken, die, obgleich sie nicht immer eine Reduktion liefern müssen, doch nach Satz 2 (für $N = 0$) immer zueinander im zweiten Sinne äquivalent sind. Infolgedessen muss X'', und daher auch X' mit einem X''', das aus $\mathfrak{C}$ durch die erwähnte Einsetzung entsteht, im zweiten Sinne äquivalent sein.

Nach Hp. 1 entsprechen $\mathfrak{A}$ und $\mathfrak{B}$ derselben Folge $\mathfrak{F}$. r sei die Ordnung, womit $\mathfrak{A}$ dem $\mathfrak{F}$ entspricht. Dann gilt $r \leqq p$. In der Tat sei angenommen, dass $r > p$ ist. Dann folgt, genau wie in Satz 1, dass $\mathfrak{A}'$ auf ein $\mathfrak{C}$ der Form (5) reduziert wird. Daher ist X'', nach dem vorigen Absatz, mit einem X''' der Form (5) im zweiten Sinne äquivalent. Dies ist aber unmöglich, weil X', und daher X'', einer mit Z anfangenden Folge lauter Variablen mit der Ordnung 0 entspricht, während X''' keiner Folge lauter Variablen mit der Ordnung 0 entsprechen kann.

Es folgt dann, wie im Satz 1, dass $\mathfrak{A}'$ und $\mathfrak{B}'$ sich auf dasselbe $\mathfrak{C}$ reduzieren. Daher sind X und Y nach dem vorletzten Absatz mit demselben X''' im zweiten Sinne äquivalent. Aber nach der Voraussetzung reduziert X' sich auf Z. Daher reduziert sich auch Y' auf Z.

Der Rest des Beweises verläuft genau wie im Satz 1.

SATZ 6. *Sind X und Y Kombinatoren, wofür*

1) *mindestens einer der beiden einer Folge von lauter Variablen entspricht,*

2) *aus den transmutativen und kommutativen Axiomen folgt, dass* $\vdash X = Y$;

dann sind X und Y äquivalent im zweiten Sinne.

Beweis: Folgt aus Satz 5, weil die betreffenden Axiome die Bedingungen der Formel (1) erfüllen.

SATZ 7. *Ax. I_2 ist nicht aus den übrigen kombinatorischen Axiomen mit Benutzung der Regeln B, C, K, W und den Eigenschaften der Gleichheit ableitbar.*

Beweis: Folgt aus Satz 6, weil die zwei Kombinatoren, die in Ax. I_2 auf den beiden Seiten des Zeichens $=$ stehen, im dritten, aber nicht im zweiten Sinne äquivalent sind.

SATZ 8. *Wenn wir in den Hypothesen von Sätzen* 1-3 *und* 5 *die folgenden Änderungen machen:*

1) $\mathfrak{A}_i$ *und* $\mathfrak{B}_i$ *brauchen nicht dieselbe Ordnung (inbezug auf ihr Entsprechen einer gemeinsamen Folge) zu haben,*

2) *nicht nur* X, *sondern auch* Y *einer Folge von lauter Variablen entspricht;*

dann folgen die Schlüsse dieser Sätze, wenn wir darin den vierten Sinn durch den dritten, und den zweiten Sinn durch den ersten ersetzen.

Beweis: Die einzigen Stellen in den Beweisen der betr. Sätze, wo wir die Voraussetzung über die Ordnung von den $\mathfrak{A}_i$ und $\mathfrak{B}_i$ benutzt haben, sind im Fall I unter den Sätzen 1 und 5, und zwar wird sie da nur benutzt, um zu beweisen, dass $\mathfrak{A}'$ und $\mathfrak{B}'$ sich auf ein gemeinsames $\mathfrak{C}$ reduzieren.

Diesen Schluss können wir auch im vorliegenden Falle erreichen. Es folgt ohne Benutzung der betr. Voraussetzung, dass entweder $\mathfrak{A}'$ und $\mathfrak{B}'$ sich auf ein gemeinsames $\mathfrak{C}$ reduzieren, oder einer der beiden auf einen Ausdruck der Form (5) reduziert wird. n sei nun so gewählt, dass nicht nur X', sondern auch Y' sich auf eine Kombination von lauter Variablen reduziert. Dies ist möglich nach Hp. 2 dieses Satzes. Dann folgt durch das Argument des dritten Absatzes des Falles 1 in den Sätzen 1 und 5, dass weder $\mathfrak{A}'$ noch $\mathfrak{B}'$ sich auf einen Ausdruck der Form (5) reduzieren lässt. Daher müssen sie sich auf einen gemeinsamen $\mathfrak{C}$ reduzieren.

Diese Änderung des n stört aber nichts in den Beweisen der betr. Sätze, ausser dass wir jetzt nicht schliessen können, dass X und Y dieselbe Ordnung haben. Also haben wir einen wirklichen Beweis, wenn wir die ganzen Beweise hindurch die Ersetzungen vom Schlusse dieses Satzes machen. Damit wird die Behauptung bewiesen.

SATZ 9. *Wenn* X *und* Y *Kombinatoren sind, wofür*

1) *sowohl* X *wie auch* Y *einer Folge von lauter Variablen entspricht,*

2) *aus den transmutativen Axiomen und Ax.* I_2 *mit Benutzung der Eigenschaften der Identität und Regeln B, C, K, W folgt, dass* $\vdash X = Y$;

dann sind X *und* Y *im dritten Sinne äquivalent.*

Beweis: Folgt aus Sätzen 3 und 8.

SATZ 10. *Die kommutativen Axiome sind nicht Folgerungen aus den anderen kombinatorischen Axiomen.*

Beweis: Die Kombinatoren, die in diesen Axiomen auf den beiden Seiten des Zeichens = stehen, sind nicht im dritten Sinne äquivalent. Daher folgt der Satz aus Satz 9.

SATZ 11. *Sind X und Y Kombinationen von Variablen und Kombinatoren derart, dass*

1) *sowohl X wie auch Y einer Folge lauter Variablen entspricht,*

2) *aus den kombinatorischen Axiomen überhaupt mit Benutzung der Eigenschaften der Gleichheit und der Regeln B, C, K, W folgt, dass $\vdash X = Y$;*

dann haben X und Y denselben Grad, und sie entsprechen derselben Folge.

Beweis: Nach den Sätzen 5 und 8 sind X und Y im ersten Sinne äquivalent. Daher folgt der Satz gleich aus der Definition der Äquivalenz.

SATZ 12. *Sind X und Y Kombinationen lauter Variablen, wofür die Hp. 2 von Satz 11 erfüllt ist, so sind X und Y identisch.*

Beweis: Nach Satz 11 entsprechen X und Y derselben Folge; dies kann nur geschehen, wenn sie Abschnitte derselben Folge sind. Weiter haben sie nach Satz 11 denselben Grad; daraus folgt, dass sie genau derselbe Abschnitt sind.

Festsetzung 9. Ein Kombinator X stellt eine Kombination Y der Variablen $x_1, x_2, \cdots, x_n$ dann und nur dann dar, wenn aus den kombinatorischen Axiomen mit Benutzung der Eigenschaften der Gleichheit und der Regeln B, C, W, K folgt, dass

$$\vdash X x_1 x_2 \cdots x_n = Y.$$

SATZ 13. *Wenn ein Kombinator eine Kombination von $x_1, x_2, \cdots, x_n$ darstellt, so stellt er nur eine dar.*

Beweis: Folgt gleich aus Satz 12.

§ 2. *Normale Kombinationen und Folgen.*

Festsetzung 1. Unter einer *normalen Kombination* von $X_0, X_1, X_2, \cdots, X_n$ verstehen wir einen Ausdruck der Form

$$(X_0 Y_1 Y_2 \cdots Y_n),$$

wo jedes Y_i eine Kombination von $X_1, X_2, \cdots, X_n$ ist.

Festsetzung 2. Hiernach wird zuweilen auch das Zeichen x_0 als Variable gebraucht.*

* In der inhaltlichen Anwendung der vorliegenden Theorie wird im allgemeinen eine Funktion (wie ϕ in II A 3), die Stelle von x_0 einnehmen. Die Variable x_0 wird hiernach im allgemeinen nur für normalen Folgen usw. benutzt.

Festsetzung 3. Unter einer *normalen Folge* (von Variablen) verstehen wir eine Folge, die durch eine normale Kombination von $x_0, x_1, x_2, \cdots, x_n$ bestimmt ist (§ 1, Festsetzung 4), wo n irgendeine ganze Zahl > 0 ist. Solche normalen Folgen werden hiernach mit griechischen Buchstaben bezeichnet.

Festsetzung 4. Unter dem *Produkt* $(\eta \cdot \zeta)$ von zwei normalen Folgen η und ζ verstehen wir die folgendermassen bestimmte Reihe (von Variablen): Es sei

$$\eta = x_0 y_1 y_2 y_3 \cdots$$
$$\zeta = x_0 z_1 z_2 z_3 \cdots.$$

Ersetzt man dann in $z_1, z_2, \cdots$ die $x_1, x_2, x_3, \cdots$ bzw. durch $y_1, y_2, y_3, \cdots$, so ist das Resultat $(\eta \cdot \zeta)$.

Satz 1. *Das Produkt von zwei normalen Folgen ist eine normale Folge.*

Beweis: η und ζ werden wie in der Festsetzung 4 bezeichnet und $(\eta \cdot \zeta)$ werde durch

$$x_0 u_1 u_2 u_3 \cdots$$

bezeichnet.

Nach der Definition einer Normalfolge gibt es ein m und ein n sodass 1) $(x_0 y_1 y_2 \cdots y_n)$ eine normale Kombination von $x_1, x_2, \cdots, x_m$ ist, die x_m wirklich enthält, 2) $y_{n+j} \equiv x_{m+j}$. In derselben Weise gibt es ein p und ein q, sodass 1) $(x_0 z_1 z_2 \cdots x_q)$ eine normale Kombination von $x_0 x_1 x_2 \cdots x_p$ ist, die weiterhin x_p wirklich enthält, und 2) $z_{q+j} \equiv x_{p+j}$. Wir können weiter annehmen, dass $p = n$ ist; denn ist $p > n$, so bleibt alles rechtig, das ich über m, n gesagt habe, wenn ich m durch $m + p - n$ ersetze, und ist $p < n$, so kann ich in ähnlicher Weise p durch n, auch q durch $q + n - p$ ersetzen.

Nach diesen Erklärungen sieht man sofort, dass u_i für $i \leqq q$ eine Kombination von $x_1, x_2, \cdots, x_m$ ist, während $u_{q+j} \equiv y_{n+j} \equiv x_{m+j}$. Daher ist $x_0 u_1 u_2 \cdots u_{q+1}$ eine normale Kombination von $x_1, x_2, \cdots, x_{m+1}$, und $x_0 u_1 u_2 u_3 \cdots$ ist die durch diese normale Kombination bestimmte Folge, w. z. b. w.

Satz 2. *Sind Y und Z Kombinatoren, die den normalen Folgen η bzw. ζ von Variablen entsprechen, so entspricht $(Y \cdot Z)$ der Folge $(\eta \cdot \zeta)$.*

Beweis: Sind η und ζ, wie in der Festsetzung 4 bezeichnet, so gibt es m, n, p, q, sodass

$$Y x_0 x_1 x_2 \cdots x_m \doteq x_0 y_1 y_2 \cdots y_n \qquad x_m \text{ nicht ausgelassen}$$
$$Z x_0 x_1 x_2 \cdots x_p \doteq x_0 z_1 z_2 \cdots z_q \qquad x_p \text{ nicht ausgelassen.}$$

Wir können ohne Beschränkung der Allgemeinheit annehmen, dass $n = p$ gilt; denn ist $p > n$, so können wir $x_{m+1}, x_{m+2}, \cdots, x_{m+p-n}$ zu den beiden Seiten der

ersten Gleichung hinzufügen, und ist $p < n$, so können wir $x_{p+1}, x_{p+2}, \cdots, x_n$ zu den beiden Seiten der zweiten Gleichung hinzufügen. Dann gilt

$$\begin{aligned}(Y \cdot Z)x_0x_1x_2 \cdots x_m &\doteq Y(Zx_0)x_1x_2 \cdots x_m \quad \text{(II B 4 Satz 1).}\\ &\doteq Zx_0y_1y_2y_3 \cdots y_n\\ &\doteq x_0u_1u_2 \cdots u_q,\end{aligned}$$

wo $u_i \equiv z_i$ mit x_i durch y_i ersetzt gilt.

§ 3. *Die Gruppierungen.*

Festsetzung 1. Eine Folge lauter Variablen heisst eine Gruppierung, wenn die Variablen darin in ihrer ursprünglichen Reihenfolge ohne Wiederholungen oder Auslassungen, aber natürlich in beliebiger Weise in Klammern zusammengefasst, erscheinen. Z. B. sind

$$x_0(x_1x_2)(x_3(x_4(x_5x_6)x_7))x_8x_9 \cdots$$
$$x_0(x_1(x_2(x_3x_4)x_5x_6x_7)x_8)x_9x_{10}$$

Gruppierungen. Jede Gruppierung ist eine normale Folge.

Festsetzung 2. Unter die Gruppierungen ist die Folge

$$x_0x_1x_2x_3 \cdots$$

einzuschliessen. Diese Gruppierung soll die *identische Gruppierung* heissen. Ihr entspricht der Identitätskombinator I.

Ich werde nun beweisen, dass jeder Gruppierung ein gewisser eindeutig bestimmter Kombinator entspricht.

Satz 1. *Der Kombinator* B_mB_n $(m \geqq 0, n > 0)$ *entspricht der Gruppierung, welche dann entsteht, wenn man* $x_{m+1}, x_{m+2}, \cdots, x_{m+n+1}$ *in einem einzigen Klammerpaar zusammenfasst. D. h.:*

$$B_mB_nx_0x_1x_2 \cdots x_{m+n+1} \doteq x_0x_1x_2 \cdots x_m(x_{m+1}x_{m+2} \cdots x_{m+n+1}).$$

Beweis:

$$\begin{aligned}B_mB_nx_0x_1x_2 \cdots x_{m+n+1} &\doteq B_n(x_0x_1x_2 \cdots x_m)x_{m+1} \cdots x_{m+n+1}\\ &\qquad \text{(vgl. II B 1, Satz 3),}\\ &\doteq x_0x_1x_2 \cdots x_m(x_{m+1}x_{m+2} \cdots x_{m+n+1}) \quad \text{(vgl. II B 1, Satz 3).}\end{aligned}$$

Satz 2. *Jeder Kombinator der Form*

$$(1) \qquad (B_{m_q}B_{n_q}) \cdot (B_{m_{q-1}}B_{n_{q-1}}) \cdot (B_{m_{q-2}}B_{n_{q-2}}) \cdots (B_{m_2}B_{n_2}) \cdot (B_{m_1}B_{n_1})$$

entspricht einer Gruppierung.

Beweis: Folgt aus Satz 1 und § 2 Satz 2, weil das Produkt (im Sinne von § 2) zweier Gruppierungen wieder eine Gruppierung ist.

SATZ 3. *Jeder Gruppierung, die nicht die identische ist, entspricht ein und nur ein Kombinator der Form* (1) *mit*

(2) $$m_q > m_{q-1} > m_{q-2} > \cdots m_2 > m_1.$$

Beweis: Wir nehmen an, dass eine Gruppierung gegeben ist, worin alle die nach I C, Def. 1 fortgeschafften Klammern, sowie auch die die gesamte Gruppierung einschliessenden, wirklich fortgeschafft sind. Die übrig bleibenden Klammern befinden sich in Paaren—eine Anfangsklammer und eine ihr zugehörige Schlussklammer—ein solches Paar nennen wir ein Klammerpaar. Wir bezeichnen dann die Gruppierung mit Γ_q, wo q die Anzahl dieser übrig bleibenden Klammerpaare ist. Es gilt $q \geqq 1$, wenn die Gruppierung nicht die identische ist.

Nun sei das Klammerpaar, dessen Anfangsklammer am weitesten links steht, als das erste angesehen. Mit diesem verknüpfen wir die Zahlen m_1, n_1 wie folgt: x_{m_1} soll das letzte x sein, das vor der Anfangsklammer steht, und $n_1 + 1$ soll die Anzahl der innerhalb des Klammerpaares stehenden Glieder sein—wo ein eingeklammerter Teilausdruck, der selbst innerhalb eines anderen Klammerpaares steht, ist als ein einziges Glied des letzteren anzusehen.

Zunächst schaffen wir das erste Klammerpaar aus Γ_q fort. Die so gestaltete Gruppierung nennen wir Γ_{q-1}. Wir suchen dann das erste Klammerpaar in Γ_{q-1}, und bestimmen davon die Zahlen m_2 und n_2 genau so wie die vorigen m_1 und n_1 aus Γ_q bestimmt wurden. Dann schaffen wir dieses Klammerpaar weg und gestalten eine neue Gruppierung Γ_{q-2}, wovon wir die Zahlen m_3 und n_3 bestimmen, u. s. w.

Nachdem wir diesen Prozess q mal wiederholt haben, kommen wir auf einer Γ_0, welche keine Klammern enthält. Dann zeige ich, dass die so konstruierten Zahlen $m_1, m_2, \cdots m_q, n_1, n_2, \cdots n_q$ die Bedingungen des Satzes erfüllen.

Zunächst ist $m_{i+1} > m_i$. Nach der Definition ist $m_{i+1} \geqq m_i$, und die Gleichheit ist unmöglich, weil wir alle die nach I C Def. 1 erlaubten Klammerauslassungen ausgeführt haben, und also zwei Anfangsklammern an derselben Stelle nicht stehen können.

Zweitens: der Kombinator (1) mit diesem m_i und n_i entspricht dem Γ_q. In der Tat sei γ_r die Gruppierung, der $B_{m_r}B_{n_r}$ nach Satz 1 entspricht, dann folgt aus der Definition der Γ_i, dass

$$\Gamma_{r+1} = \Gamma_r \cdot \gamma_{q-r} \qquad (r = 1, 2, \cdots, q-1)$$
$$\Gamma_1 = \gamma_q$$

gelten. Daher gilt (das Produkt von Folgen ist assoziativ)

$$\Gamma_q = \gamma_q \cdot \gamma_{q-1} \cdots \gamma_1.$$

Daraus folgt die Behauptung nach § 2, Satz 2.

Zuletzt gibt es nur einen Kombinator, der die Bedingungen erfüllt. Denn jeder andere Kombinator der Form (1), wofür (2) gilt, entspricht nach dem eben durchgeführten Beweis einer Gruppierung von ganz anderer Klammerstruktur. Aber derselbe Kombinator kann nicht zwei so verschiedenen Folgen entsprechen. (cf. § 1, Hilfsatz 3).

§ 4. *Die Umwandlungen.*

Festsetzung 1. Eine normale Folge von $x_0, x_1, x_2, \cdots, x_n$, worin nach den Auslassungen von I C Def. 1, keine Klammern (ausser den die gesamte Folge einschliessenden) erscheinen, nenne ich eine *Umwandlung.* (Diese Festsetzung stimmt mit der Erklärung im Abschnitte *A* überein). Z. B. sind

$$x_0x_1x_3x_1x_2x_4x_5 \cdots$$
$$x_0x_2x_4x_2x_3x_5x_6 \cdots$$

Umwandlungen, die erste ohne, die zweite mit Auslassungen.

Festsetzung 2. Die Folge:

$$x_0x_1x_2 \cdots$$

der der Kombinator *I* entspricht, ist sowohl eine Umwandlung als auch eine Gruppierung. Ich nenne sie die *identische Umwandlung.* Um weitere Umschreibungen zu vermeiden, soll hier festgestellt werden, dass diese identische Umwandlung zu allen den hierunter betrachteten Gattungen von Umwandlungen gehört.

SATZ 1. *Jede Umwandlung lässt sich in eindeutiger Weise als Produkt einer Umwandlung κ, die nur Auslassungen zulässt, wie etwa*

$$(1) \qquad (x_0x_1x_2 \cdots x_{h_1-1}\, x_{h_1+1}x_{h_1+2} \cdots x_{h_2-1}x_{h_2+1} \cdots x_{h_3-1}x_{h_3+1} \cdots x_{h_p-1}x_{h_p+1} \cdots),$$

und einer Umwandlung μ ohne Auslassungen darstellen.

Beweis: ω sei die gegebene Umwandlung. Wenn in ω keine Variablen ausgelassen werden, dann gilt $\omega = (\kappa \cdot \mu)$, wo κ die identische Umwandlung ist und $\mu = \omega$ ist. Sonst seien $x_{h_1}, x_{h_2}, \cdots, x_{h_p}$ die aus ω ausgelassenen Variablen. κ sei die Umwandlung (1), mit $h_1 \cdots h_p$ wie eben definiert. μ sei die Umwandlung, welche entsteht, wenn man in ω x_i durch x_j ersetzt, wo j aus i folgendermassen bestimmt wird: wenn $i < h_1$ ist, dann ist $j = i$; wenn $h_k < i < h_{k+1}$ ist, dann ist $j = i - k$; wenn $h_p < i$ ist, dann ist $j = i - p$. Dann ist μ eine Umwandlung ohne Auslassungen und $\omega = (\kappa \cdot \mu)$.

κ' sei nun irgendeine Umwandlung der Form (1) (bzw. die identische Umwandlung) und μ' sei eine Umwandlung ohne Auslassungen. Es sei $\omega' = (\kappa' \cdot \mu')$. Bilden wir κ'' und μ'' aus ω' genau wie wir κ und μ aus ω gebildet haben, so ist $\kappa'' = \kappa'$ und $\mu'' = \mu'$. Also wenn $\omega' = \omega$ gilt, so ist $\kappa' = \kappa$ und $\mu' = \mu$. Also sind κ und μ durch ω eindeutig bestimmt.

Satz 2. *Jedem κ, das nicht das identische ist, entspricht ein und nur ein Kombinator der Form*

(2) $$K_{h_p} \cdot K_{h_{p-1}} \cdots K_{h_2} \cdot K_{h_1},$$

wo

(3) $$h_1 < h_2 < \cdots < h_{p-1} < h_p$$

gilt.

Beweis: Es sei eine Umwandlung κ der Form (1) gegeben. Der Kombinator (2) mit dem durch (1) bestimmten $h_1, h_2, \cdots, h_p$ entspricht dann diesem κ, und die Bedingung (3) ist natürlich erfüllt. Irgendein anderer Kombinator (2), wofür (3) erfüllt ist, entspricht nach dem eben Gesagten einer von κ verschiedenen Folge κ', also nicht zu κ (§ 1, Hilfssatz 3).

Festsetzung 3. Unter einer *Permutationsfolge* verstehen wir eine normale Folge, die durch eine Permutation bestimmt ist, oder, was dasselbe ist, eine Umwandlung ohne Auslassungen oder Wiederholungen.

Satz 3. *Jede Umwandlung ohne Auslassungen lässt sich als Produkt zweier Faktoren darstellen, wovon der zweite eine Permutationsfolge ist, während im ersten die Variablen ihre ursprüngliche Reihenfolge behalten, aber wiederholt werden können. Dieser erste Faktor ist eindeutig bestimmt.*

Beweis: μ sei eine gegebene Umwandlung ohne Auslassungen. Wenn es in μ keine wiederholten Variablen gibt, so ist der zweite Faktor μ selbst, der erste die identische Umwandlung. Sonst seien $x_{k_1}, x_{k_2}, \cdots, x_{k_q}$ $(k_1 < k_2 < \cdots < k_q)$ sämtliche in μ wiederholte Variablen, und wir setzen fest, dass x_{k_1} $(r_1 + 1)$ mal, x_{k_2} $(r_2 + 1)$ mal u. s. w. bis x_{k_q} $(r_q + 1)$ mal in μ erscheinen. Dann betrachten wir die Umwandlung:

(4) $$(x_0 x_1 x_2 \cdots x_{k_1-1} x_{k_1} x_{k_1} \cdots (r_1 + 1)\,\text{mal} \cdots x_{k_1} x_{k_1+1} \cdots x_{k_2-1} x_{k_2} x_{k_2} \cdots (r_2 + 1)\,\text{mal} \cdots x_{k_2} x_{k_2+1} \cdots x_{k_q-1} x_{k_q} x_{k_q} \cdots (r_q + 1)\,\text{mal} \cdots x_{k_q} x_{k_q+1} \cdots).$$

Dann ist μ durch eine Permutation der in (4) erscheinenden Zeichen bestimmt, also ist es das Produkt von (4) und der durch diese Permutation bestimmten Permutationsfolge.

Umgekehrt sei eine Kombination (4) gegeben (wo wir unter $q = 0$ die

identische Umwandlung zu verstehen haben). Denn das Produkt von (4) nach irgendeiner Permutationsfolge ist ein μ, worin x_{k_i} $(i = 1, 2, \cdots, q)$ $(r_i + 1)$mal erscheint und kein anderes x wiederholt ist. Also kann ein μ nie zugleich als Produkt von zwei verschiedenen Kombinationen der Form (4) mit Permutationsfolgen dargestellt werden.

Der Satz wird nun bewiesen, wenn wir bemerken, dass wenn q, k_1, k_2, $\cdots, k_q$ beliebig sind, (4) die allgemeinste, den Bedingungen für den ersten Faktor genügende Folge ist.

Def. 1. $W_k{}^1 \equiv W_k \qquad k = 1, 2, 3, \cdots,$
$W_k{}^{r+1} \equiv W_k \cdot W_k{}^r \qquad k = 1, 2, 3, \cdots, \quad r = 1, 2, 3, 4, \cdots.$

SATZ 4. *Es gibt einen und nur einen Kombinator der Form*

$$W_{k_q}{}^{r_q} \cdot W_{k_{q-1}}{}^{r_{q-1}} \cdots W_{k_2}{}^{r_2} \cdot W_{k_1}{}^{r_1}, \tag{5}$$

wo ferner

$$k_1 < k_2 < \cdots < k_q \tag{6}$$

gilt, der einer gegebenen, von der identischen Umwandlung verschiedenen, den Bedingungen für den ersten Faktor im Satze 3 *genügenden Folge entspricht.**

Beweis: Zuerst: $W_k{}^r$ entspricht der Folge

$$(x_0 x_1 x_2 \cdots x_{k-1} x_k x_k x_k \cdots (r+1)\text{mal} \cdots x_k x_{k+1} x_{k+2} \cdots).$$

In der Tat für $r = 1$ folgt dies aus II B 3, Satz 4. Ist es für ein gegebenes r angenommen, dann haben wir für $r + 1$

$$\begin{aligned} W_k{}^{r+1} x_0 x_1 x_2 \cdots x_k &\doteq W_k{}^r x_0 x_1 \cdots x_{k-1} x_k x_k \\ &\doteq x_0 x_1 x_2 \cdots x_{k-1} x_k x_k \cdots (r+2)\text{mal} \cdots x_k. \end{aligned}$$

Es wird nun bewiesen werden, dass, wenn (6) gilt, (5) wie es geschrieben steht dem Ausdruck (4) entspricht. Zu diesem Behuf kürzen wir (5) mit q durch s ersetzt mit $\mathfrak{W}_s$, und den Ausdruck

$$\begin{aligned} &(x_0 x_1 \cdots x_{k_1-1} x_{k_1} x_{k_1} \cdots (r_1 + 1)\text{mal} \\ &\quad \cdots x_{k_1} x_{k_1+1} \cdots x_{k_s-1} x_{k_s} x_{k_s} \cdots (r_s + 1)\text{mal} \cdots x_{k_s}) \end{aligned}$$

mit X_s ab. Dann haben wir schon für $s = 1$ bewiesen,

$$\mathfrak{W}_s x_0 x_1 x_2 \cdots x_{k_s} \doteq X_s \tag{7}$$

Ist dies für ein bestimmtes s vorausgesetzt, so haben wir

---

* Dieser Satz und Lemma 3 meiner oben zit. Abhandlung sind wesentlich äquivalent. Der hier gegebene Beweis ist alternativ zu jenem.

$$
\begin{aligned}
\mathfrak{W}_{s+1}x_0x_1x_2\cdots x_{k_{s+1}} &\doteq W_{k_{s+1}}{}^{r_{s+1}}(\mathfrak{W}_s x_0)x_1x_2\cdots x_{k_{s+1}} \\
&\doteq \mathfrak{W}_s x_0x_1x_2\cdots x_{k_{s+1}-1}x_{k_{s+1}}x_{k_{s+1}}\cdots(r_{s+1})\text{mal}\cdots x_{k_{s+1}} \\
&\qquad\qquad \text{(nach dem eben bewiesenen)}, \\
&\doteq X_s x_{k_s+1}x_{k_s+2}\cdots x_{k_{s+1}-1}x_{k_{s+1}}x_{k_{s+1}}\cdots(r_{s+1}+1)\text{mal}\cdots x_{k_{s+1}} \\
&\qquad\qquad \text{(nach der Voraussetzung)}, \\
&\doteq X_{s+1} \qquad\qquad \text{(nach der Festsetzung über } X_s\text{)}.
\end{aligned}
$$

Also wird durch Induktion (7) für $s = q$, also die Behauptung bewiesen.

Der Beweis des Satzes folgt gleich. Denn wenn wir die Konstanten $q, k_1, k_2, \cdots, k_q$ in (5) einsetzen, so entspricht der resultierende Kombinator der Folge (4) nach dem letzten Absatz. Wenn wir andere Konstanten, die (6) genügen, in (5) einsetzen, so entspricht der resultierende Kombinator einer ganz anderen Folge der Form (4). Also gibt es nur einen Kombinator der betreffenden Beschaffenheit.

SATZ 5. *Jeder Permutationsfolge entspricht ein Kombinator $\mathfrak{C}$, der aus einem Produkt lauter $C_1, C_2, \cdots$ besteht, und zwar so, dass das mit dem höchsten Index versehene C_n nur einmal vorkommt.*

Beweis: Nach einem wohlbekannten Satz über Permutationen ist jede Permutation der Elemente $x_1, x_2, \cdots, x_m$ ein Produkt von Transformationen benachbarter Elementen. Dies bedeutet, in unsere Terminologie übersetzt, dass jede Permutationsfolge, die durch eine Permutation von $x_1, x_2, \cdots, x_m$ bestimmt wird, ein Produkt der Folgen

$$
\begin{gathered}
(x_0x_2x_1x_3\cdots) \\
(x_0x_1x_3x_2\cdots) \\
\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot \\
(x_0x_1x_2\cdots x_{m-2}x_mx_{m-1}x_{m+1}\cdots)
\end{gathered}
$$

ist. Diesen Folgen entsprechen bzw. die Kombinatoren $C_1, C_2, \cdots, C_{m-1}$. Infolgedessen entspricht der gegebenen Permutationsfolge ein $\mathfrak{C}$, das aus einem Produkt von lauter $C_1, C_2, \cdots, C_{m-1}$ besteht (§ 2, Satz 2).

Es sei nun eine Permutationsfolge π gegeben, die x_{m+1}, aber keine mit höherem Index versehene Variable, wirklich permutiert. x_k sei die Variable, die die Stelle von x_{m+1} einnimmt. Dann ist π ein Produkt von zwei Folgen π_1, und π_2, wo

$$\pi_1 = (x_0x_1x_2\cdots x_{k-1}x_{k+1}\cdots x_{m-1}x_mx_{m+1}x_kx_{m+2}\cdots)$$

ist und π_2 durch eine Permutation von $x_1x_2\cdots x_m$ bestimmt ist. Der Folge π_1 entspricht aber der Kombinator $\mathfrak{C}_1$,

$$\mathfrak{C}_1 \equiv C_k \cdot C_{k+1} \cdots C_{m-1} \cdot C_m.$$

Der Folge π_2 entspricht weiter nach dem vorigen Absatz ein $\mathfrak{C}_2$, das ein Produkt lauter $C_1, C_2, \cdots C_{m-1}$ ist. Also entspricht $\mathfrak{C} = \mathfrak{C}_1 \cdot \mathfrak{C}_2$ der Folge π, und $\mathfrak{C}$ erfüllt die Bedingungen des Satzes, weil C_m nur einmal vorkommt.

§ 5. *Darstellung der allgemeinen normalen Folge.*

SATZ 1. *Jede normale Folge lässt sich in eindeutiger Weise als Produkt einer Umwandlung und einer Gruppierung darstellen.*

Beweis: η sei die gegebene Folge. Wir erzeugen aus η eine Umwandlung ω und eine Gruppierung γ folgendermassen: zuerst schaffen wir alle die innerhalb η erscheinenden Klammern fort, dann soll der resultierende Ausdruck ω heissen. Zweitens lassen wir in η die Klammern stehen und schaffen die Variablen fort, und füllen dann die Leerstellen, wo Variablen früher waren, mit $x_0, x_1, x_2, \cdots$ von links nach rechts in ihrer naturgemässen Reihenfolge, ohne Auslassungen oder Wiederholungen aus. Der neue Ausdruck ist eine Gruppierung, γ. Diese ω und γ nennen wir die mit η assoziierte Umwandlung bzw. Gruppierung. Nach der Festsetzung 4, § 2 gilt $\eta = (\omega \cdot \gamma)$.

Nun sei ω' irgendeine Umwandlung und γ eine Gruppierung. Es sei $\eta' = (\omega' \cdot \gamma')$. Dann sind die mit η' assoziierte Gruppierung bzw. Umwandlung genau dieses ω' bzw. γ'. Infolgedessen muss, wenn $\eta' = \eta$ ist, auch $\omega' = \omega$ und $\gamma' = \gamma$ sein.

SATZ 2. *Jeder normalen Folge entspricht mindestens ein Kombinator der Form:*

$$(\mathfrak{K} \cdot \mathfrak{W} \cdot \mathfrak{C} \cdot \mathfrak{B}),$$

wo a) $\mathfrak{K}$ *in der Form von* § 4 *Satz* 2 *steht,*
b) $\mathfrak{W}$ *in der Form von* § 4 *Satz* 4 *steht,*
c) $\mathfrak{C}$ *in der Form von* § 4 *Satz* 5 *steht,*
d) $\mathfrak{B}$ *in der Form von* § 3 *Satz* 3 *steht.*

Ferner sind $\mathfrak{K}$, $\mathfrak{W}$ *und* $\mathfrak{B}$ *durch diese Bedingungen eindeutig bestimmt.*

Beweis: η sei eine gegebene normale Folge. Nach Satz 1 und § 4, Sätzen 1, 3 gibt es eine Gruppierung γ, eine Umwandlung κ, die nur Auslassungen zulässt, eine Umwandlung ω, die nur Wiederholungen zulässt, und eine Permutationsfolge, π, derart, dass

$$\eta = (\kappa \cdot \omega \cdot \pi \cdot \gamma)$$

ist. Nach § 3 Satz 3, § 4 Sätze 2, 4, 5 gibt es gewisse die Bedingungen a-d erfüllende $\mathfrak{K}$, $\mathfrak{W}$, $\mathfrak{C}$, $\mathfrak{B}$ die diesen κ bzw. ω bzw. π bzw. γ entsprechen. Dann entspricht $(\mathfrak{K} \cdot \mathfrak{W} \cdot \mathfrak{C} \cdot \mathfrak{B})$ der Folge η nach § 2 Satz 2.

Nun sei irgendeine der Folge η entsprechende Kombinator der Form (1) gegeben, etwa $(\mathfrak{K}' \cdot \mathfrak{W}' \cdot \mathfrak{C}' \cdot \mathfrak{B}')$. Es seien κ', ω', π', γ', die zu $\mathfrak{K}'$, $\mathfrak{W}'$, $\mathfrak{C}'$ bzw. $\mathfrak{B}'$ gehörenden Folgen; sie gehören denselben Kombinationsgattungen wie κ, ω, π bzw. γ an. Der betreffende Kombinator entspricht dann $(\kappa' \cdot \omega' \cdot \pi' \cdot \gamma')$ (§ 2, Satz 2); also, wenn er auch dem η entspricht, gilt

$$(\kappa' \cdot \omega' \cdot \pi' \cdot \gamma') = (\kappa \cdot \omega \cdot \pi \cdot \gamma) \qquad \text{(§ 1 Hilfsatz 3).}$$

Also gelten $\gamma' = \gamma$ und $(\kappa' \cdot \omega' \cdot \pi) = (\kappa \cdot \omega \cdot \pi')$ (Satz 1); $\kappa' = \kappa$ und $(\omega' \cdot \pi') = (\omega \cdot \pi)$ (§ 4 Satz 1); $\omega' = \omega$ (§ 4 Satz 3). Daher sind $\mathfrak{B}'$ und $\mathfrak{B}$ identisch (§ 3 Satz 3); $\mathfrak{K}'$ und $\mathfrak{K}$ sind identisch (§ 2 Satz 2); $\mathfrak{W}'$ und $\mathfrak{W}$ sind identisch (§ 4 Satz 4). Also sind $\mathfrak{K}$, $\mathfrak{W}$, und $\mathfrak{B}$ durch die Bedingungen eindeutig bestimmt.

D. Reguläre Kombinatoren.

§ 1. *Vorläufige Festsetzungen und Sätze.*

Festsetzung 1. Ein Kombinator X heisst *regulär,* wenn er die Form

$$(X_1 \cdot X_2 \cdot \cdots \cdot X_n)$$

hat, wo jedes X_i ferner von einer der Formen

$$B_p B_q,\ C_q,\ W_q,\ K_q,\ B_p I,$$

ist. Die einzelne X_i heissen die *Glieder* von X.

Festsetzung 2. Ein Kombinator heisst *normal,* bzw. *in der normalen Form,* wenn er in der in II C 5 Satz 2 besprochenen Form steht.

Festsetzung 3. Ein Kombinator X heisst in einer gegebenen Form *umformbar,* wenn es ein schon in der betreffenden Form stehendes X' gibt, sodass $\vdash X = X'$.

In diesem Abschnitte beweise ich den Hauptsatz: wenn immer zwei reguläre Kombinatoren X und Y derselben Folge entsprechen, dann $\vdash X = Y$. Dies folgt daraus, dass erstens jeder reguläre Kombinator sich in die Normalform umformen lässt, und zweitens der Hauptsatz gilt, wenn nur X und Y normal sind. Der zweite in Abschnitt A erwähnte Hauptsatz wird hier für normale Kombinationen bewiesen.

Festsetzung 4. Zum Zwecke der Abkürzung möchte ich die folgenden Buchstaben für gewisse Gattungen regulärer Kombinatoren gebrauchen, derart, dass besondere Kombinatoren dadurch bezeichnet werden, dass Indizes an das betreffende Gattungszeichen angeheftet werden.

$\mathfrak{B}$: sämtliche Glieder der Form $B_m B_n$ oder $B_m I$.
$\mathfrak{C}$: sämtliche Glieder der Form C_n oder $B_m I$
$\mathfrak{K}$: sämtliche Glieder der Form K_n oder $B_m I$
$\mathfrak{M}$: sämtliche Glieder der Form W_n, C_n oder $B_m I$
$\mathfrak{W}$: sämtliche Glieder der Form W_n oder $B_m I$
Ω: sämtliche Glieder der Form C_n, K_n, W_n, oder $B_m I$.

SATZ 1. *Zu jedem regulären Kombinator X gibt es ein X' derart, dass 1) X' regulär ist, 2) X' wenn von I selbst verschieden, gar keine Glieder der Form $B_m I$ enthält, während die anderen Glieder genau dieselben wie in X sind, und* 3) $\vdash X = X'$.

Beweis: Klar aus II B 2, Satz 1 und B 4, Satz 4. Wenn sämtliche Glieder in X der Form $B_m I$ sind, so ist X' gleich I, sonst ist X' von I verschieden.

SATZ 2. $\vdash BB \cdot C_1 = C_1 \cdot C_2 \cdot B$.

Beweis: Wir haben zunächst aus Ax. $(CC)_1$ und II B 2, Satz 1

$$\vdash C_1 \cdot C_1 = B_2 I = I$$
$$\vdash C_2 \cdot C_2 = B(C_1 \cdot C_1) \qquad \text{(II B 4, Satz 2; II B 3, Def. 1).}$$
$$= B_3 I = I. \qquad (\text{Ax. } (CC)_1).$$

Daher:
$$\vdash BB \cdot C_1 = C_1 \cdot C_2 \cdot C_2 \cdot C_1 \cdot BB \cdot C_1$$
$$= C_1 \cdot C_2 \cdot B \cdot C_1 \cdot C_1 \qquad (\text{Ax. } (BC)),$$
$$= C_1 \cdot C_2 \cdot B. \qquad \text{w. z. b. w.}$$

§ 2. *Die kommutativen Gesetze.*

Festsetzung 1. Eine Gleichung der Form

$$\vdash C_1 B_{m+1} X = BX \cdot B_n$$

heisst ein *kommutatives Gesetz* für X, weil es eine gewisse Art von Vertauschbarkeit von X mit anderen Etwasen gewährt. Einige der Axiome sind von dieser Form; diese habe ich kommutative Axiome genannt.

SATZ 1. *Wenn X ein Etwas ist, wofür*

$$\vdash CB_{m+1}X = BX \cdot B_n;$$

dann gilt für irgendein Etwas Y

$$\vdash B_m Y \cdot X = X \cdot B_n Y.$$

Beweis: Aus der Hp. und I D Satz 3 folgt

(1)	$\vdash CB_{m+1}XY = (BX \cdot B_n)Y.$	
Aber	$\vdash CB_{m+1}XY = B_{m+1}YX$	(Reg. C),
	$= (B(B_mY))X$	(II B 1, Satz 5),
(2)	$= B_mY \cdot X$	(II B 4, Def. 1).
Auch	$\vdash (BX \cdot B_n)Y = BX(B_nY)$	(II B 4, Satz 2),
(3)	$= X \cdot B_nY.$	

aus (1), (2), (3) wird der Satz bewiesen.

SATZ 2. *Wenn X ein Etwas ist, wofür $\vdash CB_{m+1}X = BX \cdot B_n$; dann*

$$\vdash CB_{m+k+1}X = BX \cdot B_{n+k} \qquad (k = 1, 2, 3, \cdots).$$

Beweis: Wir haben zunächst mit Anwendungen der Eigenschaften der Gleichheit und Definitionen,

$$\vdash (CB_{m+1}X) \cdot B_k = (BX \cdot B_n) \cdot B_k$$

(1)	$= BX \cdot B_{n+k}$	(II B 4, Sätze 3 und 5).
aber	$\vdash CB_{m+1}X \cdot B_k = B(CB_{m+1}X)B_k$	(II B 4, Def. 1),
	$= BB(CB_{m+1})XB_k$	(Reg. B),
(2)	$= (BB \cdot C)B_{m+1}XB_n$	(II B 4, Satz 1),
	$= (C_1 \cdot C_2 \cdot B)B_{m+1}XB_n$	(§ 1, Satz 2),
	$= C_1(C_2(BB_{m+1}))XB_n$	(II B 4, Sätze 1 und 3),
	$= C_2(BB_{m+1})\, B_nX$	(Reg. C),
	$= C_1(BB_{m+1}B_n)X$	(Def. von C_2; II B 3, Def. 1; Reg.
(3)	$= CB_{m+k+1}X$	(II B 4, Def. 1 und Satz 5).

Aus (1) und (3) wird der Satz bewiesen.

SATZ 3. *Wenn X ein Etwas ist, wofür $\vdash CB_{m+1}X = BX \cdot B_n$; dann gilt für ein beliebiges Etwas Y*

(1) $$\vdash B_{m+k+h}Y \cdot B_hX = B_hX \cdot B_{n+k+h}Y \quad (k, h = 0, 1, 2 \cdots).$$

Beweis: Nach Satz 2 haben wir

$$\vdash CB_{m+k+1}X = BX \cdot B_{n+k}.$$

$\therefore$ nach Satz 1, $\vdash B_{m+k}Y \cdot X = X \cdot B_{n+k}Y.$

Die Behauptung (1) folgt dann aus II B 4, Satz 6 und II B 1, Satz 5.

SATZ 4. *Wenn X ein beliebiges Etwas ist; dann gilt für $m = 0, 1, 2, \cdots$, $n = 1, 2, \cdots$, und $p \geqq m + 1$,*

$$\vdash B_{p+n}Y \cdot B_mB_n = B_mB_n \cdot B_pY.$$

Beweis: Für $n = 1$ folgt der Satz aus Ax. B und Satz 3. Ist der Satz für ein gegebenes n angenommen, so wird er folgendermassen für $n + 1$ bewiesen:

$$\begin{aligned} \vdash B_{p+n+1}Y \cdot B_m B_{n+1} &= B_{p+n+1}Y \cdot B_m B_n \cdot B_m B && \text{(II B 4, Satz 7)},\\ &= B_m B_n \cdot B_{p+1}Y \cdot B_m B && \text{(Voraussetzung)},\\ &= B_m B_n \cdot B_m B \cdot B_p Y && \text{(dieser Satz für } n = 1),\\ &= B_m B_{n+1} \cdot B_p Y && \text{(II B 4, Satz 7)}. \end{aligned}$$

SATZ 5. *Wenn X ein beliebiges Etwas ist; dann sind die folgenden Gleichungen beweisbar,*

a) $\vdash B_m Y \cdot C_p = C_p \cdot B_m Y,$ *wenn* $m \geqq p \geqq 1$ *gilt,*
b) $\vdash B_m Y \cdot W_p = W_p \cdot B_{m+1} Y,$ *wenn* $m \geqq p \geqq 1$ *gilt,*
c) $\vdash B_m Y \cdot K_p = K_p \cdot B_{m-1} Y,$ *wenn* $m \geqq p \geqq 1$ *gilt.*

Beweis: Diese Gleichungen folgen aus Satz 3, den Axiomen C, W und K und den Definitionen von II B 3.

SATZ 6. *Das Axiom I_1 lässt sich aus den übrigen kombinatorischen Axiomen beweisen.*

Beweis:

$$\begin{aligned} \vdash CBI &= CB(WK) && \text{(I C, Def. 3)},\\ &= B(CB)WK && \text{(Reg. } B),\\ &= (B \cdot C)BWK && \text{(II B 4, Satz 1)},\\ &= (C_2 \cdot C_1 \cdot BB)BWK && \text{(Ax. } (BC)),\\ &= C_2(C_1(BBB))WK && \text{(II B 4, Satz 1)},\\ &= C_1(C_1 B_2 W)K && \text{(Def. von } C_2;\ \text{Reg. } B;\ \text{II B 1, Def. 1)},\\ &= C_1(B_2 W B_2)K && \text{(Ax. } W;\ \text{II B 4, Satz 1; II B 1, Satz 5)},\\ &= B_2 C_1 B_2 W B_2 K && \text{(II B 1, Satz 3)},\\ &= C_1 B_3 C_1 W B_2 K && \text{(Ax. } C;\ \text{II B 4, Satz 1; II B 1, Satz 5)},\\ &= B_3 W C_1 B_2 K && \text{(Reg. } C),\\ &= BW(C_1 B_2 K) && \text{(II B 1, Satz 2)},\\ &= BW(BK \cdot I) && \text{(Ax. } K),\\ &= W \cdot K_2 \cdot I && \text{(II B 4, Def. 1; Def. von } K_2),\\ &= W \cdot C_1 \cdot K \cdot I && \text{(Ax. } (CK)),\\ &= W \cdot K \cdot I && \text{(Ax. } (WC)),\\ &= BI \cdot I, \quad \text{w. z. b. w.} && \text{(Ax. } (WK)). \end{aligned}$$

§ 3. *Umformung in die Form $\Omega \cdot \mathfrak{B}$.*

SATZ 1. *Jedes $\dot{\mathfrak{B}}$ kann entweder in I oder in die Normalform von II C 3, Satz 3, nämlich*

(1) $$B_{m_q}B_{n_q}\cdot B_{m_{q-1}}B_{n_{q-1}}\cdot\cdot\cdot B_{m_2}B_{n_2}\cdot B_{m_1}B_{n_1},$$
wo
(2) $$m_1 < m_2 < \cdot\cdot\cdot < m_q$$
gilt, umgeformt werden.

Beweis: Nach Festsetzung 4 und § 1, Satz 1 kann jedes $\mathfrak{B}$ entweder in I oder in die Form (1) umgeformt werden. Es bleibt nur zu beweisen, dass das betreffende $\mathfrak{B}$ im letzten Fall so umgeformt werden kann, dass auch (2) gilt. Ich beschränke mich auf solche $\mathfrak{B}$s.

Aus § 2, Satz 4 und II B 4, Satz 7 haben wir

(3) $$\vdash B_mB_n\cdot B_pB_q = B_{n+p}B_q\cdot B_mB_n, \quad \text{wenn } p > m \text{ gilt.}$$
(4) $$\vdash B_mB_n\cdot B_pB_q = B_mB_{n+q}, \quad \text{wenn } p = m \text{ gilt.}$$

Nun sei $\mathfrak{B}$ schon in der Normalform, dann kann $B_rB_s\cdot\mathfrak{B}$ in die Normalform umgeformt werden. In der Tat sei

$$r < m_t, m_{t+1}, \cdot\cdot\cdot, m_q \text{ und entweder } t = 1, \text{ oder } m_{t-1} \leqq r$$

Dann ist nach (3)

$$B_rB_s\cdot\mathfrak{B} = B_{m_q+s}B_{n_q}\cdot B_{m_{q-1}+s}B_{n_{q-1}}\cdot\cdot\cdot B_{m_t+s}B_{n_t}\cdot B_rB_s\cdot B_{m_{t-1}}B_{n_{t-1}}\cdot\cdot\cdot B_{m_1}B_{n_1},$$

wo natürlich, wenn $t = 1$ gilt, die Glieder rechts von B_rB_s an der rechten Seite nicht da sind. Wenn $t = 1$ oder $r > m_{t-1}$ gilt, ist die rechte Seite der eben Geschrieben schon in der Normalform. Sonst kann B_rB_s mit seinem rechtsstehenden Nachbarn nach (4) verschmolzen werden, und der neue Ausdruck wird in der Normalform sein.

Nun sei $\mathfrak{B}$ ein beliebiger Ausdruck der Form (1). $\mathfrak{B}_r$ sei ($r = 1, 2, \cdot\cdot\cdot$) das Produkt der r rechtsstehenden Glieder von $\mathfrak{B}$. $\mathfrak{B}_1$ ist schon in der Normalform. Wenn $\mathfrak{B}_r$ in die Normalform umgeformt werden kann, so kann $\mathfrak{B}_{r+1} \equiv B_{m_{r+1}}B_{n_{r+1}}\cdot\mathfrak{B}_r$ nach dem vorigen Absatz in die Normalform umgeformt werden. Also kann $\mathfrak{B}_q \equiv \mathfrak{B}$ in die Normalform umgeformt werden.

Satz 2. *Zu jedem $\mathfrak{B}$ und C_p gibt es ein $\mathfrak{B}'$ und ein $\mathfrak{C}'$ derart, dass*

$$\vdash \mathfrak{B}\cdot C_p = \mathfrak{C}'\cdot\mathfrak{B}'.$$

Beweis: Für $\mathfrak{B} \equiv I$, klar.

Zunächst sei $\mathfrak{B} \equiv B_mB$. Dann unterscheiden wir vier Fälle:

***Fall* 1:** $p > m + 1$. Dann

$$\vdash B_mB\cdot C_p = C_{p+1}\cdot B_mB \qquad (\S\ 2, \text{Satz } 4).$$

***Fall* 2:** $p = m + 1$. Dann

$$\begin{aligned} \vdash B_m B \cdot C_{m+1} &= B_m(B \cdot C_1) && \text{(II B 3, Satz 1; II B 4, Satz 6)}, \\ &= B_m(C_2 \cdot C_1 \cdot BB) && \text{(Ax. } (BC)\text{)}, \\ &= C_{m+2} \cdot C_{m+1} \cdot B_{m+1}B && \text{(II B 3, Satz 1; II B 1, Satz 5; II B 4, Satz 6)}. \end{aligned}$$

Fall 3: $p = m > 0$. Dann gilt

$$\begin{aligned} \vdash B_m B \cdot C_m &= B_{m-1}(BB \cdot C) \\ &= B_{m-1}(C_1 \cdot C_2 \cdot B) && \text{(§ 1, Satz 2)}, \\ &= C_m \cdot C_{m+1} \cdot B_{m+1}B && \text{(II B 3, Satz 1; II B 4, Satz 6)}. \end{aligned}$$

Fall 4: $p < m$. Dann gilt

$$\vdash B_m B \cdot C_p = C_p \cdot B_m B \qquad \text{(§ 2, Satz 5a)}.$$

Also ist der Satz für $\mathfrak{B} \equiv B_m B$ bewiesen. Es folgt durch Induktion, dass es zu einem beliebigen $\mathfrak{C}$ ein $\mathfrak{C}'$ und ein $\mathfrak{B}'$ derart gibt, dass

$$\vdash B_m B \cdot \mathfrak{C} = \mathfrak{C}' \cdot \mathfrak{B}'.$$

Das allgemeinste $\mathfrak{B}$ kann nun entweder in I oder in ein Produkt von N Faktoren der Form $B_m B$ (Satz 1, II B 4, Satz 7)* umgeformt werden. Wir können nun ohne Beschränkung der Allgemeinheit $\mathfrak{B}$ als in dieser letzten Form gegeben betrachten. $\mathfrak{B}_M$ sei das Produkt der M rechtsstehenden Faktoren von $\mathfrak{B}$. Wenn der Satz für jedes $\mathfrak{B}_M$ bewiesen ist, so ist er für $\mathfrak{B}$ bewiesen. Aber für $M = 1$ ist er schon im letzten Absatz bewiesen. Für ein bestimmtes M sei angenommen, dass $\vdash \mathfrak{B}_M \cdot C_p = \mathfrak{C} \cdot \mathfrak{B}_M'$, dann gilt

$$\begin{aligned} \vdash \mathfrak{B}_{M+1} \cdot C_p &= B_m B \cdot \mathfrak{B}_M \cdot C_p \\ &= B_m B \cdot \mathfrak{C} \cdot \mathfrak{B}_M' \\ &= \mathfrak{C}' \cdot \mathfrak{B}'' \cdot \mathfrak{B}_M' \\ &= \mathfrak{C}' \cdot \mathfrak{B}'. \end{aligned}$$

Daher folgt der Satz durch Induktion für alle $\mathfrak{B}_M$, also auch für $\mathfrak{B}$.

SATZ 3. *Wenn X ein regulärer Kombinator ist, dessen sämtliche Glieder der form $B_m B_n$ oder C_p sind, so kann X in die Form $(\mathfrak{C} \cdot \mathfrak{B})$ umgeformt werden.*

Beweis: Es sei $X \equiv X_1 \cdot X_2 \cdot \cdot \cdot X_n$ wo die X_i die Glieder von X sind.

Es sei nun angenommen, dass $(X_1 \cdot X_2 \cdot \cdot \cdot X_q)$ in die Form $(\mathfrak{C}' \cdot \mathfrak{B}')$ umgeformt werden kann; dann gilt, wenn $X_{q+1} \equiv B_m B_n$,

* Für das $\mathfrak{B}$ von II C 3 (1) Satz 2 ist $N = n_1 + n_2 + n_3 + \cdot \cdot \cdot + n_q$.

$$\begin{aligned}\vdash X_1 \cdot X_2 \cdot X_3 \cdots X_q \cdot X_{q+1} &= \mathfrak{C}' \cdot \mathfrak{B}' \cdot B_m B_n \\ &= \mathfrak{C}' \cdot \mathfrak{B}'',\end{aligned}$$

während, wenn $X_{q+1} \equiv C_p$ ist, gilt

$$\begin{aligned}\vdash X_1 \cdot X_2 \cdots X_{q+1} &= \mathfrak{C}' \cdot \mathfrak{B}' \cdot C_p \\ &= \mathfrak{C}' \cdot \mathfrak{C}'' \cdot \mathfrak{B}'' \qquad \text{(Satz 2)}, \\ &= \mathfrak{C}''' \cdot \mathfrak{B}''.\end{aligned}$$

Also ist der Satz durch Induktion auf q für X bewiesen, weil er für $q = 1$ klar ist.

SATZ 4. *Für jedes* $\mathfrak{B}$ *und* W_p *gibt es ein* Ω *und ein* $\mathfrak{B}'$ *derart, dass*

$$\vdash \mathfrak{B} \cdot W_p = \Omega \cdot \mathfrak{B}'.$$

Beweis: Für $\mathfrak{B} \equiv I$ klar. Ich beschränke mich also auf den Fall $\mathfrak{B} \not\equiv I$. Zunächst zeige ich, dass es für jedes $m = 0, 1, 2, \cdots$, $p = 0, 1, 2, \cdots$, $k = 0, 1, 2, \cdots, p-1$, ein $q > 0$, ein $h < q$ und ein X, deren sämtliche Glieder der Form $B_m B_n$ oder C_p sind, derart gibt, dass

$$\vdash B_m B \cdot W_p \cdot W_{p-1} \cdots W_{p-k} = W_q \cdot W_{q-1} \cdots W_{q-h} \cdot X. \tag{7}$$

Es sind drei Fälle zu unterscheiden:

Fall 1. $p \leqq m$. Weil nach § 2, Satz 5 für alle $r \leqq m$

$$\vdash B_m B \cdot W_r = W_r \cdot B_{m+1} B,$$

so haben wir hier,

$$\vdash B_m B \cdot W_p \cdot W_{p-1} \cdots W_{p-k} = W_p \cdot W_{p-1} \cdots W_{p-k} \cdot B_{m+k+1} B.$$

Fall 2. $p = m + 1$. Hier gilt für $k = 0$,

$$\begin{aligned}\vdash B_m B \cdot W_{m+1} &= B_m (B \cdot W) && \text{(II B 4, Satz 6)}, \\ &= B_m (W_2 \cdot W_1 \cdot C_2 \cdot B_2 B \cdot B) && \text{(Ax. } (BW)), \\ &= W_{m+2} \cdot W_{m+1} \cdot C_{m+2} \cdot B_{m+2} B \cdot B_m B && \text{(II B 4, Satz 6)}\end{aligned}$$

Für $k \geqq 1$,

$$\begin{aligned}\vdash B_m B \cdot W_{m+1} \cdot W_m \cdots W_{m-k+1} & \\ = W_{m+2} \cdot W_{m+1} \cdot C_{m+2} \cdot B_{m+2} B \cdot B_m B \cdot W_m \cdots W_{m-k+1} & \quad \text{(nach dem Falle } k = 0), \\ = W_{m+2} \cdot W_{m+1} \cdot C_{m+2} \cdot W_m \cdot W_{m-1} \cdots W_{m-k+1} \cdot B_{m+k+2} B \cdot B_{m+k} B & \quad \text{(nach Fall 1)}, \\ = W_{m+2} \cdot W_{m+1} \cdot W_m \cdots W_{m-k+1} \cdot C_{m+k+2} \cdot B_{m+k+2} B \cdot B_{m+k} B, &\end{aligned}$$

wobei der letzte Schritt aus § 2, Satz 5 und Definition von C_{m+k+2} folgt.

Fall 3. $p > m + 1$. Aus § 2, Satz 4 folgt, für $k = 0$,

$$\vdash B_mB \cdot W_p = W_{r+1} \cdot B_mB.$$

Also:

$$\begin{aligned}\vdash B_mB \cdot W_p \cdot W_{p-1} \cdots W_{m+3} \cdot W_{m+2} \cdot W_{m+1} \cdots W_{p-k},\\ = W_{p+1} \cdot W_p \cdots W_{m+3} \cdot W_{m+2} \cdot B_mB \cdot W_{m+1} \cdots W_{p-k},\\ = W_{p+1} \cdot W_p \cdots W_{m+3} \cdot W_{m+2} \cdots W_{p-k} \cdot C_{m+k+2} \cdot B_{m+k+2}B \cdot B_{m+k}B\end{aligned}$$

(Fall 2).

Also ist (7) bewiesen. Durch den Induktionsprozess, den ich im letzten Absatz des Beweises von Satz 2 benutzt habe, wird die Gleichung bewiesen, die entsteht wenn man in (7) B_mB durch ein beliebiges $\mathfrak{B}$ ersetzt. Wenn man in dieser Gleichung $k = 0$ setzt, so hat man

$$\vdash \mathfrak{B} \cdot W_p = W_q \cdot W_{q-1} \cdots W_{q-h} \cdot X,$$

wo, nach Satz 3 $\quad \vdash X = \mathfrak{C} \cdot \mathfrak{B}'.$

Also $\quad \vdash \mathfrak{B} \cdot W_p = W_q \cdot W_{q-1} \cdots W_{q-h} \cdot \mathfrak{C} \cdot \mathfrak{B}'$
$\quad = \Omega \cdot \mathfrak{B}' \qquad$ w. z. b. w.,

weil $(W_q \cdot W_{q-1} \cdots W_{q-h} \cdot \mathfrak{C})$ die Definition eines Ω erfüllt.

Satz 5. *Für jedes* $\mathfrak{B}$ *und* K_p *gibt es ein* $\mathfrak{K}$ *und ein* $\mathfrak{B}'$ *derart, dass* $\vdash \mathfrak{B} \cdot K_p = \mathfrak{K} \cdot \mathfrak{B}'$ *oder auch, gibt es ein* $\mathfrak{K}$, *wofür* $\vdash \mathfrak{B} \cdot K_p = \mathfrak{K}$.

Beweis: Ich beweise den Satz zunächst für den Fall $\mathfrak{B} \equiv B_mB$. Dann gibt es drei Fälle:

Fall 1: $p \leqq m$. Dann gilt nach § 2, Satz 5

$$\vdash B_mB \cdot K_p = K_p \cdot B_{m-1}B.$$

Fall 2: $p = m + 1$. Dann folgt aus Ax. (BK) und II B 4, Satz 6

$$\vdash B_mB \cdot K_{m+1} = B_m(B \cdot K) = B_m(K_1 \cdot K_1) = K_{m+1} \cdot K_{m+1}.$$

Fall 3: $p > m + 1$. Dann folgt aus § 2, Satz 4,

$$\vdash B_mB \cdot K_p = K_{p+1} \cdot B_mB.$$

Der Rest des Beweises läuft genau wie in Satz 2.

Satz 6. *Jeder reguläre Kombinator lässt sich in die Form* $(\Omega \cdot \mathfrak{B})$ *umformen.*

Beweis: X sei ein regulärer Kombinator und $X_1, X_2, \cdots, X_q$ seien seine Glieder, so dass

$$X \equiv X_1 \cdot X_2 \cdots X_n.$$

Der Satz ist sicher wahr für X_1. Nehmen wir an, er ist für den Kombinator $(X_1 \cdot X_2 \cdot \cdot \cdot X_q)$ wahr, dann werde ich ihn für $(X_1 \cdot X_2 \cdot \cdot \cdot X_{q+1})$ beweisen. In der Tat sei

$$\vdash X_1 \cdot X_2 \cdot \cdot \cdot X_q = \Omega' \cdot \mathfrak{B}'.$$

Dann ist X_{q+1} entweder $B_m B_n$, C_p, W_p oder K_p.* Im ersten Fall ist das zu Beweisende klar, wenn wir $\Omega \equiv \Omega'$, $\mathfrak{B} \equiv \mathfrak{B}' \cdot B_m B_n$ setzen. In anderen Fällen wissen wir aus den Sätzen 2, 4 und 5, dass es ein Ω'' und ein $\mathfrak{B}''$ gibt, wofür $\vdash \mathfrak{B}' \cdot X_{q+1} = \Omega'' \cdot \mathfrak{B}''$ gilt, also

$$\begin{aligned} \vdash X_1 \cdot X_2 \cdot \cdot \cdot X_{q+1} &= \Omega' \cdot \mathfrak{B}' \cdot X_{q+1} \\ &= \Omega' \cdot \Omega'' \cdot \mathfrak{B}'' \\ &= \Omega \cdot \mathfrak{B}, \end{aligned}$$

wenn wir $\Omega \equiv \Omega' \cdot \Omega''$, $\mathfrak{B} \equiv \mathfrak{B}''$ definieren.

§ 4. *Die Umformung* $\Omega = \mathfrak{K} \cdot \mathfrak{M}$.

SATZ 1. *Jedes* Ω *kann in die Form* $(\mathfrak{K} \cdot \mathfrak{M})$ *umgeformt werden.*

Beweis: Es genügt zu zeigen, dass jeder Kombinator der Form $(\mathfrak{M} \cdot K_p)$ in die betreffende Form übergeführt werden kann, denn das allgemeinste Ω enthält entweder kein K—und dann ist der Satz klar $(\mathfrak{K} \equiv I)$—, oder es kann in die Form

$$(\mathfrak{M}_1 \cdot K_{p_1} \cdot \mathfrak{M}_2 \cdot K_{p_2} \cdot \cdot \cdot \mathfrak{M}_k \cdot K_{p_k} \cdot \mathfrak{M}_{k+1})$$

umgeformt werden, wo einzelne $\mathfrak{M}_i \equiv I$ können. (In der Tat folgt dies durch Einschaltungen von gewissen I's, welche durch II B 4, Satz 4 erlaubt sind). Dann wird durch Wiederholung des Prozesses wodurch $(\mathfrak{M} \cdot K_p)$ in die Form des Satzes umgeformt wird, der ganze Ausdruck in diese Form gebracht.

Weiter genügt es zu beweisen, dass $W_m \cdot K_p$ in die Form $K_r \cdot W_s$ bzw. I und $C_m \cdot K_p$ in die Form $K_r \cdot C_s$ bzw. K_r umgeformt werden können. Denn wenn diese Behauptungen bewiesen sind, so folgt daraus, dass die einzelnen Glieder eines $\mathfrak{M}$ eins nach dem andern über die K's übertragen oder mit ihnen verschmolzen werden können.

Die Behandlung von $(C_m \cdot K_p)$ gibt vier Fälle:

Fall 1: $p \leqq m - 1$. Dann

$$\vdash C_m \cdot K_p = K_p \cdot C_{m-1} \qquad (§ 2, \text{Satz 5c; II B 3}).$$

* Wir können natürlich annehmen, dass keine Glieder der Form $B_m I$ vorkommen, weil der Satz für $X \equiv I$ klar ist (s. § 1, Satz 1).

Fall 2: $p = m$. Dann

$$\vdash C_m \cdot K_m = B_{m-1}(C_1 \cdot K_1) \quad \text{(II B 4, Satz 6; II B 3)},$$
$$= B_{m-1}K_2 \quad \text{(Ax. ((CK)))},$$
$$= K_{m+1} \quad \text{(II B 3, Satz 5)}.$$

Fall 3: $p = m + 1$. Nach Fall 2 folgt

$$\vdash C_m \cdot K_{m+1} = C_m \cdot C_m \cdot K_m$$
$$= B_m(C_1 \cdot C_1) \cdot K_m \quad \text{(II B 3, Satz 1)},$$
$$= B_{m+2}I \cdot K_m \quad \text{(Ax. } (CC)_1\text{; II B 1, Satz 5)},$$
$$= K_m \quad \text{(II B 2, Satz 1, und II B 4, Satz 4)}.$$

Fall 4: $p > m + 1$. Dann nach § 2, Satz 5,

$$\vdash C_m \cdot K_p = K_p \cdot C_m.$$

Die Behandlung für $(W_m \cdot K_p)$ gibt drei Fälle:

Fall 1: $p \leqq m - 1$.

$$\vdash W_m \cdot K_p = K_p \cdot W_{m-1} \quad \text{(§ 2, Satz 5c; II B 3, Sätze 3 u. 5)}.$$

Fall 2: $p = m$.

$$\vdash W_m \cdot K_m = B_{m-1}(W_1 \cdot K_1) \quad \text{(II B 4, Satz 6; II B 3, Sätze 3 u. 5)},$$
$$= B_m I \quad \text{(Ax. } (WK)\text{; II B 1, Satz 5)},$$
$$= I \quad \text{(II B 2, Satz 1)}.$$

Fall 3: $p > m$. Dann

$$\vdash W_m \cdot K_p = K_{p-1} \cdot W_m \quad \text{(§ 2, Satz 5b; II B 3, Sätze 3 u. 5)}.$$

Damit ist der Satz vollständig bewiesen.

SATZ 2. *Jedes* $\mathfrak{K}$ *kann entweder in I oder in die Normalform von II C 4, Satz 3, nämlich*

(1) $$(K_{h_p} \cdot K_{h_{p-1}} \cdots K_{h_2} \cdot K_{h_1})$$

wo

(2) $$h_1 < h_2 < \cdots < h_p$$

sind, umgeformt werden.

Beweis: Nach § 1, Festsetzung 4 und § 1, Satz 1 kann $\mathfrak{K}$ entweder auf I oder auf die Form (1) gebracht werden. Aus § 2, Satz 5 folgt

(3) $$\vdash K_m \cdot K_p = K_{p+1} \cdot K_m, \text{ wenn } p \geqq m.$$

Wenn es in dem betreffenden Ausdruck zwei benachbarte K's etwa K_{h_s} und $K_{h_{s-1}}$ gibt, wofür $h_{s-1} \geqq h_s$ ist, so kann eine gewisse Vertauschung stattfinden.

Nach einer gewissen Anzahl von Vertauschungen nach (3) wird der Ausdruck auf eine Form, wo (2) zutrifft, gebracht. Der genaue Beweis verläuft hier wie im § 3, Satz 1.

§ 5. *Die Normalform für $\mathfrak{M}$.*

In meiner oben erwähnten Abhandlung habe ich schon bewiesen, dass aus gewissen Axiomen (besser Axiomenschemen, wovon einige unendlich viele Axiome enthalten) die folgenden sich schliessen lassen: 1) jedes $\mathfrak{M}$ kann in die Normalform umgeformt werden, 2) wenn $\mathfrak{M}_1$ und $\mathfrak{M}_2$ derselben Folge entsprechen, so folgt $\vdash \mathfrak{M}_1 = \mathfrak{M}_2$. Um diese Ergebnisse unserer Theorie zu sichern, genügt es zu beweisen, dass die dort gegebenen Axiomen, und auch die Definitionen von $W_2, W_3 \cdots$ aus unserem Grundgerüst ableitbar sind.

Satz 1. $\vdash C_m \cdot C_m = I \qquad (m = 1, 2, 3 \cdots).$

Beweis: Nach Definition von C_m und II B 4, Satz 6 gilt

$$\begin{aligned} \vdash C_m \cdot C_m &= B_{m-1}(C_1 \cdot C_1) \\ &= B_{m-1}(B_2 I) && (\text{Ax. } (CC)_1), \\ &= I && (\text{II B 1, Satz 5; II B 2, Satz 1}). \end{aligned}$$

Satz 2. $\vdash C_m \cdot C_{m+1} \cdot C_m = C_{m+1} \cdot C_m \cdot C_{m+1} \qquad (m = 1, 2 \cdots).$

Beweis: $$\begin{aligned} \vdash C_m \cdot C_{m+1} \cdot C_m &= B_{m-1}(C_1 \cdot C_2 \cdot C_1) && (\text{II B 4, Sätze 3 und 6}), \\ &= B_{m-1}(C_2 \cdot C_1 \cdot C_2) && (\text{Ax. } (CC)_2), \\ &= C_{m+1} \cdot C_m \cdot C_{m+1}. \end{aligned}$$

Satz 3. $\vdash C_m \cdot C_{m+j} = C_{m+j} \cdot C_m,$ *wenn* $j > 1,$ $(m = 1, 2 \cdots).$

Beweis: Folgt aus § 2, Satz 5, wenn wir C_{j-1} für Y in die Gleichung a) setzen.

Satz 4. $\vdash C_m \cdot W = W \cdot C_{m+1} \qquad (m = 2, 3, 4, \cdots).$

Beweis: Folgt gleich aus § 2, Satz 5b.

Satz 5. $\vdash W_m \cdot W_n = W_n \cdot W_{m+1} \qquad (m \geqq n = 1, 2, 3, \cdots).$

Beweis: Für $m = n$,

$$\begin{aligned} \vdash W_m \cdot W_m &= B_{m-1}(W_1 \cdot W_1) && (\text{II B 3, Def. 2; II B 4, Satz 6}), \\ &= B_{m-1}(W_1 \cdot W_2) && (\text{Ax. } (WW)), \\ &= W_m \cdot W_{m+1} && (\text{II B 3, Def. 2; II B 4, Satz 6}). \end{aligned}$$

Für $m > n$ folgt der Satz aus § 2, Satz 5b.

SATZ 6. $\vdash W_{m+1} = C_m \cdot W_m \cdot C_{m+1} \cdot C_m$ $\quad (m = 1, 2, 3 \cdots)$.

Beweis: $\vdash C_m \cdot W_m = B_{m-1}(C_1 \cdot W_1)$ (II B 3, Def. 2; II B 4, Satz 6).
$= B_{m-1}(W_2 \cdot C_1 \cdot C_2)$ (Ax. (CW)),
$= W_{m+1} \cdot C_m \cdot C_{m+1}$ (II B 3, Def. 2; II B 4, Satz 6).

also $\vdash W_{m+1} = (W_{m+1} \cdot C_m \cdot C_{m+1}) \cdot C_{m+1} \cdot C_m$ (Satz 1),
$= C_m \cdot W_m \cdot C_{m+1} \cdot C_m$ w. z. b. w.

SATZ 7. *Wenn $\mathfrak{M}_1$ und $\mathfrak{M}_2$ derselben Folge lauter Variablen entsprechen, dann* $\vdash \mathfrak{M}_1 = \mathfrak{M}_2$.

Beweis: In meiner oben zitierten Abhandlung gegeben. Die Voraussetzungen jenes Beweises sind in der Tat schon hier bewiesen, wie folgt:

dort	hier
Axiomschema I	Satz 1
" II	Satz 2
" III	Satz 3
" IV	Satz 4
" V	Ax. (WC)
" VI	Folgen aus Satz 5 durch Umkehrung des Beweises der
" VII	Gleichungen (6) und (7) meiner zitierten Abhandlung.
Definition von W_k	Satz 6.

Jener Beweis lässt sich aber vermöge der hier vorliegenden Entwicklungen bedeutend abkürzen. In der Tat können wir aus § 2, Satz 5b und Ax. (CW) in einer den Beweisen von § 3, Sätzen 2 und 4 ähnlicher Weise schliessen, dass ein $\mathfrak{M}$ in die Form $\mathfrak{W} \cdot \mathfrak{C}$ umgeformt werden kann, und dann weiter, wie im § 3, Satz 1 nachweisen, dass $\mathfrak{W}$ sich in die Normalform umformen lässt. Dabei werden Lemmas 1 und 2 jener Abhandlung bewiesen. Für Lemmas 3 und 4 sind alternative Beweise schon in II C 4, Sätzen 4 und 5 geliefert.

SATZ 8. *Jedes $\mathfrak{M}$ lässt sich in die Normalform umformen.*

Beweis: Dies ist im Laufe des Beweises von Satz 7 dargetan. (Lemmas 1 und 2 meiner früheren Abhandlung).—Der Satz folgt auch direkt aus Satz 7, § 6 (unten) Satz 2, und II C 5, Satz 2.

§ 6. *Zusammenfassung und Schluss.*

SATZ 1. *Jeder reguläre Kombinator kann in die Normalform umgeformt werden.*

Beweis: Jeder reguläre Kombinator X lässt sich in die Form $(\Omega \cdot \mathfrak{B})$, wo $\mathfrak{B}$ in der Normalform steht, umformen (§ 3, Sätze 1 und 6). Dieses Ω lässt sich in die Form $(\mathfrak{K} \cdot \mathfrak{M})$ umformen, wo $\mathfrak{K}$ in der Normalform ist (§ 4, Sätze 1 und 2). Endlich lässt sich $\mathfrak{M}$ in die Normalform umformen (§ 4, Satz 8). Also kann X in die Normalform $(\mathfrak{K} \cdot \mathfrak{W} \cdot \mathfrak{C} \cdot \mathfrak{B})$ umgeformt werden.

Satz 2. *Jeder reguläre Kombinator entspricht einer normalen Folge lauter Variablen, und zwar im ersten Sinne.*

Beweis: Die einzelnen Glieder eines regulären Kombinator entsprechen solchen Folgen (II B 3; II C 3, Satz 1; II B 2). Daher entspricht das Produkt einer solchen Folge (II C 2, Satz 2). Dass er der Folge im ersten Sinne entspricht, ist aus dem Beweis von II C 2, Satz 2 ohne weiteres ersichtlich.

Satz 3. *Wenn X_1 und X_2 reguläre Kombinatoren sind, wofür $\vdash X_1 = X_2$; dann sind X_1 und X_2 äquivalent in dem dritten Sinne.*

Beweis: Nach II C 1, Satz 11, und Satz 2 sind X_1 und X_2 im ersten Sinne äquivalent, also entsprechen sie beide einer gemeinsamen Folge lauter Variablen. Nach Satz 2 entsprechen sie dieser Folge im ersten Sinne. Also ist der Sinn der Äquivalenz zwischen X_1 und X_2 der dritte.

Satz 4. *Wenn X_1 und X_2 reguläre, derselben Folge von lauter Variablen entsprechende Kombinatoren sind; dann $\vdash X_1 = X_2$.*

Beweis: Sind Y_1 und Y_2 reguläre, in der Normalform stehende Kombinatoren, in welche X_1 bzw. X_2 umgeformt werden können (Satz 1), so entsprechen Y_1 und Y_2 derselben Folge wie X_1 und X_2 (Hp. und Satz 3).

Sind	$Y_1 \equiv \mathfrak{K}_1 \cdot \mathfrak{M}_1 \cdot \mathfrak{B}_1$ und $Y_2 \equiv \mathfrak{K}_2 \cdot \mathfrak{M}_2 \cdot \mathfrak{B}_2$,		
dann	$\vdash \mathfrak{B}_1 = \mathfrak{B}$ und $\vdash \mathfrak{K}_1 = \mathfrak{K}_2$		(II C 5, Satz 2),
und	$\vdash \mathfrak{M}_1 = \mathfrak{M}_2$		(§ 5, Satz 7).
Daher	$\vdash Y_1 = Y_2$.		
also	$\vdash X_1 = X_2$,	w. z. b. w.	

Satz 5. *Damit für zwei reguläre Kombinatoren X_1 und X_2 $\vdash X_1 = X_2$ gilt, ist es notwendig und hinreichend, dass X_1 und X_2 im dritten Sinne äquivalent sind.*

Beweis: Klar aus Sätzen 3 und 4.

Festsetzung 1. Eine normale Folge ξ hat die Ordnung n, wenn 1) es

eine Kombination X von $x_0, x_1, x_2, \cdots, x_n$ gibt, sodass die Folge durch (Xx_{n+1}) bestimmt ist, und 2) n die keinste Zahl ist, wofür ein solches X existiert.

Es folgt aus dieser Festsetzung, dass jeder Kombinator der dem ξ entspricht, ihm mindestens mit der Ordnung $n+1$ entspricht.*

SATZ 6. *ξ sei eine normale Folge der Ordnung n, und X sei ein der Folge ξ entsprechender normaler Kombinator. Dann entspricht X der Folge ξ mit der Ordnung $n+1$.*

Beweis: Wir nehmen ein m so gross, dass X', wobei

$$X' \equiv Xx_0x_1x_2\cdots x_m,$$

sich auf einen Abschnitt von ξ reduziert. Wenn in dieser Reduktion keine der Variablen $x_{n+1}, x_{n+2}, \cdots, x_m$ gestört werden, so ist der Satz bewiesen. Sonst führen wir die Reduktion von X' ohne Störung von $x_{n+1}, x_{n+2}, \cdots, x_m$ soweit fort, bis wir auf einen Ausdruck der Form

$$Y(Zx_0)y_1y_2\cdots y_q$$

kommen (wo Y ein Glied von X_k, Z † ein Produkt solcher Glieder ist, und $y_1\cdots y_q$ Kombinationen von $x_1x_2\cdots x_m$ sind), sodass eine weitere Reduktion auf einen Ausdruck derselben Form ohne Störung von $x_{n+1}\cdots x_m$ nicht möglich ist. Wir unterscheiden dann vier Fälle:

1) Y ist ein K_p. Dann wird ein x_s, $s > n$, in der weiteren Reduktion ausgelassen. Weil durch Reduktionsprozesse keine Variablen eingesetzt werden, so bleibt x_s ausgelassen bis zur Ende der Reduktion von X'. Weil dieses x_s nicht in ξ ausgelassen ist, kann X nicht der Folge ξ entsprechen.

2) Y ist ein W_p. Dann wird in der weiteren Reduktion ein x_s, $s > n$, verdoppelt. Weil X normal ist, so kann kein Glied der Form K_p in Z vorkommen; also bleibt x_s verdoppelt bis zur Ende. Weil x_s nicht in ξ verdoppelt ist, so kann X auch in diesem Falle nicht der Folge ξ entsprechen.

3) Y ist ein C_p. In diesem Falle führen wir die Reduktion fort, bis wir an einen Ausdruck der obigen Form ankommen, wo nun Y das C_p mit höchstem Index ist. Durch dieses C_p wird ein höchstes x_s $s > n$ mit einer niedrigeren x_t vertauscht, und weil dieses C_p nur einmal vorkommt (§ 1, Fest-

---

* Wir haben hier $n+1$, nicht n, weil ich die Variable x_0 zugelassen habe. Die Behauptung folgt, weil in jeder Reduktion auf einen Abschnitt von ξ die Variable x_n gestört werden muss.

† Streng genommen, können wir statt (Zx_0) einen Ausdruck haben, worauf (Zx_0) sich reduziert; aber dies stört den Kern des Beweises gar nicht.

setzung 2), so kann x_s nie seine Stelle wieder erreichen. Aber dies widerspricht noch einmal der Voraussetzung, dass X der Folge ξ entspricht.

4) Y ist ein B_pB_q. Dann reduziert sich X' auf eine Kombination, worin mindestens ein x_s, $s > n$, eingeklammert ist. Daher entspricht X nicht der Folge ξ.

Diese vier Fälle erschöpfen alle möglichkeiten, weil Glieder der Form B_mI in einem normalen Kombinator nicht vorkommen.

SATZ 7. *Wenn X eine beliebige normale Kombination von lauter Variablen ist, so gibt es einen normalen Kombinator, der sie darstellt.*

Beweis: Wir nehmen an, dass X eine normale Kombination der Variablen $x_0, x_1, \cdots, x_n$ ist. Y sei der normale Kombinator, welcher der durch X bestimmten Folge entspricht (II C 5, Satz 2). Die Ordnung dieses Entsprechens ist $\leqq n+1$ (Satz 6, Festsetzung 1). Also muss $(Yx_0x_1x_2 \cdots x_n)$ sich aus X reduzieren, und daher wird ipso facto X durch Y dargestellt.

E. EIGENTLICHE KOMBINATOREN.

§ 1. *Vorläufige Festsetzungen und Sätze.*

Festsetzung 1. Ein Kombinator heisst *eigentlich,* wenn er einer Folge lauter Variablen entspricht.

In diesem Abschnitte beweise ich, dass jeder eigentliche Kombinator in der Form $\mathfrak{R}I$, wo $\mathfrak{R}$ regulär ist, umgeformt werden kann. Daraus folgt, hinsichtlich der Ergebnisse des letzten Abschnitts, dass zwei derselben Folge entsprechende Kombinatoren immer gleich sind. Der Beweis der in Abschnitt A erwähnten Hauptsätze II und III wird hier vollzogen (der letzte für eigentliche Kombinatoren).

Festsetzung 2. Ausser den Gattungszeichen von II D 1, Festsetzung 4 benutze ich den Buchstaben $\mathfrak{R}$ für einen regulären Kombinator.

Festsetzung 3. Ein Kombinator heist *regulierbar,* wenn er in einen regulären Kombinator umgeformt werden kann; d. h. wenn es einen regulären Kombinator gibt, der ihm gleich ist.

SATZ 1. *Sind die Kombinatoren X und Y regulierbar, so ist auch $(X \cdot Y)$ regulierbar.*

Beweis: Nach den Voraussetzungen gibt es $\mathfrak{R}_1$ und $\mathfrak{R}_2$, sodass $\vdash X = \mathfrak{R}_1$ und $\vdash Y = \mathfrak{R}_2$, also $\vdash X \cdot Y = \mathfrak{R}_1 \cdot \mathfrak{R}_2$. $(\mathfrak{R}_1 \cdot \mathfrak{R}_2)$ ist aber regulär (dies folgt direkt aus II D 1, Festsetzung 1).

SATZ 2. *Ist der Kombinator X regulierbar, so ist jedes $(\mathfrak{B}X)$ regulierbar.*

Beweis: Wenn $\mathfrak{B} \equiv I$ ist, klar.

Zunächst sei $\mathfrak{B} \equiv B_m$. Setzen wir dann

$$\vdash X = \mathfrak{N}, \quad \mathfrak{N} \equiv X_1 \cdot X \cdot \cdots X_n.$$

Dann $\vdash B_m X = B_m X_1 \cdot B_m X_2 \cdots B_m X_n$ (II B 4, Sätze 3 u. 6),

und die rechte Seite ist regulär.

Es sei nun ein allgemeines $\mathfrak{B}$ gegeben. Wir können annehmen, dass $\mathfrak{B}$ in der Normalform steht. Dann folgt wenn $m_1 = 0$ ist (wo m_1 wie in II C 3, Satz 3 zu verstehen ist),

$$\vdash \mathfrak{B} = B\mathfrak{B}' \cdot B_{n_1}$$

also $$\vdash \mathfrak{B}X = B\mathfrak{B}'(B_{n_1}X) = \mathfrak{B}' \cdot B_{n_1}X.$$

Die rechte Seite ist regulierbar nach dem eben Bewiesenen und Satz 1 Dagegen sei $m_1 > 0$. Dann

$$\vdash \mathfrak{B}X = B_{m_1}\mathfrak{B}'X = B(B_{m_1-1}\mathfrak{B}')X$$
$$= B_{m_1-1}\mathfrak{B}' \cdot X.$$

Die rechte Seite ist wieder regulierbar nach dem oben Gesagten und Satz 1.

SATZ 3. *Wenn X und Y beliebige Etwase sind, dann* $\vdash XY = (X \cdot BY)I$.

Beweis: Klar aus II B 2, Satz 4, und II B 4, Satz 1.

SATZ 4. *Jeder Kombinator der Form $(\mathfrak{N}I)$ entspricht einer Folge lauter Variablen, und zwar in dem ersten Sinne.*

Beweis: n sei so gewählt, dass der Ausdruck $(\mathfrak{N}x_0x_1x_2 \cdots x_n)$ sich auf eine normale Kombination von $x_0, x_1, x_2, \cdots, x_n$, etwa $(x_0y_1y_2 \cdots y_q)$ ohne Auslassung von x_n reduziert (möglich nach II D 6, Satz 2). Dann wird $(\mathfrak{N}Ix_1x_2 \cdots x_n)$ auf $(Iy_1y_2 \cdots y_q)$ im ersten Sinne reduziert. Dass sich die weitere Reduktion auf $(y_1y_2 \cdots y_q)$ im ersten Sinne vollzieht, ist selbstverständlich. Also entspricht $(\mathfrak{N}I)$ der durch die eben geschilderte Kombination bestimmte Folge.

SATZ 5. *Eine notwendige und hinreichende Bedingung dafür, dass ein $(\mathfrak{N}I)$ einer normalen Folge entspricht, ist, dass es ein $\mathfrak{N}'$ und ein $\mathfrak{B}$ gibt, sodass*

$$\vdash \mathfrak{N} = B\mathfrak{N}' \cdot \mathfrak{B}.$$

Beweis: Die Bedingung ist hinreichend; denn ist sie erfüllt, so gilt

$$\vdash \mathfrak{N}I = (B\mathfrak{N}' \cdot \mathfrak{B})I = \mathfrak{N}' \cdot \mathfrak{B}I.$$

$(\mathfrak{B}I)$ ist regulierbar nach Satz 2; also ist $(\mathfrak{N}I)$ regulierbar nach Satz 1. Daher entspricht $(\mathfrak{N}I)$ einer normalen Folge (Satz 4; II D 6, Satz 2; II C 1, Satz 11).

Die Bedingung ist notwendig. In der Tat sei angenommen, dass $(\mathfrak{N}Ix_1x_2\cdots x_n)$ sich auf eine normale Kombination V von $x_1, x_2, \cdots, x_n$ reduziert. Dann erscheint x_1 in V vereinzelt und an der ersten Stelle. $\mathfrak{N}$ werde in die Normalform umgeformt, etwa

$$\vdash \mathfrak{N} = \mathfrak{K}\cdot\mathfrak{W}\cdot\mathfrak{C}\cdot\mathfrak{B}.$$

Dann ist $\mathfrak{K}$ von der Faktor K_1 frei, weil sonst x_1 in V ausfallen würde, also

$$\vdash \mathfrak{K} = B\mathfrak{K}'.$$

Gleichfalls ist $\mathfrak{W}$ von der Faktor W_1 frei, weil sonst x_1 in V verdoppelt sein würde, also $\vdash \mathfrak{W} = B\mathfrak{W}'$. Weiter entspricht $\mathfrak{C}$ einer durch eine Permutation der Variablen $x_2, x_3, \cdots, x_m$ bestimmten Folge, also ist $\mathfrak{C}$ in ein Produkt von $C_2, C_3, \cdots, C_{n-1}$, umformbar * und daher $\vdash \mathfrak{C} = B\mathfrak{C}'$. Aus den letzten drei Formeln folgt

$$\vdash \mathfrak{N} = B(\mathfrak{K}'\cdot\mathfrak{W}'\cdot\mathfrak{C}')\cdot\mathfrak{B} \qquad \text{w. z. b. w.}$$

SATZ 6. *Zu jeder Folge lauter Variablen gibt es ein $\mathfrak{N}_1$, und zwar ein normales $\mathfrak{N}_1$ ohne Glieder der Form B_n, sodass $(\mathfrak{N}_1I)$ der Folge entspricht. Gibt es überdies ein anderes der Folge entsprechendes $\mathfrak{N}_2$, so gilt für ein durch $\mathfrak{N}_2$ bestimmtes n*

$$\vdash \mathfrak{N}_2 = \mathfrak{N}_1\cdot B_n.$$

Beweis: Wir nehmen an, die Variablen in der gegebenen Folge sind $x_1, x_2, x_3, \cdots$. Die Folge sei etwa

$$(1) \qquad x_jy_1y_2y_3\cdots \qquad j \geqq 1.$$

wo y_i eine Kombination gewissen x's ist. $\mathfrak{N}_1$ sei ein normaler Kombinator, welcher der Folge

$$(2) \qquad x_0x_jy_1y_2y_3\cdots$$

entspricht. Dann entspricht $(\mathfrak{N}_1I)$ der gegebenen Folge nach dem Beweis von Satz 4. Enthält $\mathfrak{N}_1$ ein Glied der Form B_n, so müsste $\mathfrak{N}_1$, weil es normal ist, von der Form $(\mathfrak{N}_1'\cdot B_n)$ sein; aber in diesem Falle würde $\mathfrak{N}_1$ einer Folge entsprechen, worin eine Anfangsklammer links von der zweiten Variablen steht. Weil (2) diese Form nicht hat, so erfüllt $\mathfrak{N}_1$ die Bedingungen des ersten Teils des Satzes.

---

* Vgl. Beweis von II C 4, Satz 5.

Nun sei $\mathfrak{R}_2$ irgendein regulärer Kombinator derart, dass $(\mathfrak{R}_2 I)$ der gegebenen Folge entspricht. Wir können ohne Beschränkung der Allgemeinheit annehmen, dass $\mathfrak{R}_2$ normal ist (II D 6, Satz 1). Wenn $\mathfrak{R}_2$ Glieder der Form B_n enthält, so gibt es ein $\mathfrak{R}_2'$ ohne solche Glieder, und ein B_n, sodass

$$\vdash \mathfrak{R}_2 = \mathfrak{R}_2' \cdot B_n.$$

Dann gilt $\vdash \mathfrak{R}_2 I = \mathfrak{R}_2'(B_n I) = \mathfrak{R}_2' I$ (II B 2, Satz 1; II B 2, Satz 1). Im entgegengesetzten Falle setzen wir $\mathfrak{R}_2' \equiv \mathfrak{R}_2$. $\mathfrak{R}_2'$ entspricht in den beiden Fällen einer Folge der Form

$$x_0 x_k z_1 z_2 z_3 \cdots,$$

(d. h. ohne Klammern vor der zweiten Variable.) Daher entspricht $\mathfrak{R}_2 I$ nach dem Beweis von Satz 4 der Folge:

$$x_k z_1 z_2 z_3 \cdots.$$

Weil dies mit der gegebenen Folge übereinstimmen muss, so ist $k = j$, $z_1 = y_1$, $z_2 = y_2$ u. s. w. $\mathfrak{R}_2'$ entspricht daher derselben Folge wie $\mathfrak{R}_1$. Also:

$$\vdash \mathfrak{R}_2' = \mathfrak{R}_1 \qquad \text{(II D 6, Satz 4).}$$

$$\therefore \quad \vdash \mathfrak{R}_2 = \mathfrak{R}_1 \cdot B_n \qquad \text{w. z. b. w.}$$

Festsetzung 4. Eine von der Variablen x_0 frei Folge ξ heisst der Ordnung n, wenn 1) es eine Kombination X von $x_1, x_2, \cdots x_n$ gibt, sodass die Folge durch $X x_{n+1}$ bestimmt wird, 2) n die kleinste Zahl dieser Beschaffenheit ist.

SATZ 7. *Dass* $(\mathfrak{R}_1 I)$ *von Satz* 6 *entspricht seiner Folge mit einer Ordnung, die mit der Ordnung der Folge selbst übereinstimmt.*

Beweis: Das $\mathfrak{R}_1$ entspricht seiner normalen Folge mit der Ordnung $n+1$, wo n die Ordnung der Folge selbst ist. (II D 6, Satz 6). Wie im Satz 4 folgt daraus, dass $(\mathfrak{R}_1 I)$ seiner Folge mit der Ordnung n entspricht.

SATZ 8. *Zu jeder Kombination lauter Variablen gibt es mindestens einen Kombinator, der sie darstellt.*

Beweis: Folgt aus Sätzen 6 und 7.

§ 2. *Die Kombinatoren* Γ *und eine Verallgemeinerung der kommutativen Gesetze.* Diese Sätze sind Hilfssätze für § 3 unten.

Def. 1. $\Gamma_1 \equiv C_1$; $\Gamma_{n+1} \equiv \Gamma_n \cdot C_{n+1}$, $(n = 1, 2, 3, \cdots)$.

SATZ 1. $\vdash \Gamma_n = C_1 \cdot C_2 \cdots C_n$.

Beweis: Klar.

Satz 2. *Wenn $X_0, X_1, \cdots, X_n, Y$ beliebige Etwase sind, so gilt* $\vdash \Gamma_n X_0 Y X_1 X_2 \cdots X_n = X_0 X_1 \cdots X_n Y$.

Beweis: Für $n = 1$, klar aus Regel C.

Ist nun der Satz für ein bestimmtes n angenommen, dann wird er für $n + 1$ wie folgt bewiesen:

$$\begin{aligned} \vdash \Gamma_{n+1} X_0 Y X_1 \cdots X_{n+1} &= \Gamma_n (C_{n+1} X_0) X_1 Y X_2 \cdots X_n && \text{(Def. 1; II B 4, Satz 1)} \\ &= C_{n+1} X_0 X_1 \cdots X_n Y X_{n+1} && \text{(Voraussetzung),} \\ &= X_0 X_1 \cdots X_n X_{n+1} Y && \text{(II B 3, Satz 2).} \end{aligned}$$

Also folgt der Satz durch Induktion.

Satz 3. $\vdash \Gamma_{n+1} = C_1 \cdot B\Gamma_n$.

Beweis: Für $n = 1$ klar.

Ist der Satz für ein bestimmtes n angenommen, so gilt für dieses n

$$\begin{aligned} \vdash \Gamma_{n+2} &= \Gamma_{n+1} \cdot C_{n+2} && \text{(Def. 1),} \\ &= C_1 \cdot B\Gamma_n \cdot C_{n+2} && \text{(Hp.),} \\ &= C_1 \cdot B(\Gamma_n \cdot C_{n+1}) && \text{(II B 3, Def. 1; II B 4, Sätze 2, 3),} \\ &= C_1 \cdot B\Gamma_{n+1} && \text{(Def. 1).} \end{aligned}$$

Also folgt der Satz durch Induktion.

Satz 4. $\vdash BB \cdot \Gamma_n = \Gamma_{n+1} \cdot B$.

Beweis: Für $n = 1$ ist dies in II D 1, Satz 2 bewiesen.

Ist der Satz für ein bestimmtes n angenommen, dann

$$\begin{aligned} \vdash BB \cdot \Gamma_{n+1} &= BB \cdot \Gamma_n \cdot C_{n+1} && \text{(Def. 1),} \\ &= \Gamma_{n+1} \cdot B \cdot C_{n+1} && \text{(Hp.),} \\ &= \Gamma_{n+1} \cdot C_{n+2} \cdot B && \text{(II D 2, Satz 4),} \\ &= \Gamma_{n+2} \cdot B && \text{(Def. 1).} \end{aligned}$$

Also wird der Satz durch Induktion bewiesen.

Satz 5. *Wenn Y ein beliebiges Etwas ist; dann*

$$\vdash B_p(C_1 B_{n+1} Y) = \Gamma_{p+1}(B_p B_{n+1}) Y, \qquad (p = 0, 1, 2, 3 \cdots).$$

Beweis: Definieren wir vorübergehend

$$X_p \equiv \Gamma_p (B_{p-1} B_{n+1}) Y, \qquad (p = 1, 2, 3 \cdots);$$

dann $\quad X_1 \equiv C_1 B_{n+1} Y,$

und

$$\begin{aligned}\vdash BX_p &= B_2B\Gamma_p(B_{p-1}B_{n+1})Y && \text{(II B 1, Satz 3),}\\ &=(BB\cdot\Gamma_p)(B_{p-1}B_{n+1})Y && \text{(II B 1, Satz 5; II B 4, Def. 1),}\\ &=\Gamma_{p+1}(B(B_{p-1}B_{n+1}))Y && \text{(Satz 4; II B 4, Satz 1),}\\ &=\Gamma_{p+1}(B_pB_{n+1})X && \text{(II B 1, Satz 5),}\\ &=X_{p+1}.\end{aligned}$$

Also folgt der Satz aus II B 1, Satz 4.

SATZ 6. *Wenn X, Y beliebige Etwase sind, so gilt*

$$\vdash \Gamma_{p+1}(B_pB_{n+1}\cdot X)Y=\Gamma_{p+2}(B_{p+1}B_{n+1})YX.$$

Beweis:

$$\begin{aligned}\vdash \Gamma_{p+1}(B_pB_{n+1}\cdot X)Y &= \Gamma_{p+1}(B_{p+1}B_{n+1}X)Y && \text{(II B 4, Def. 1; II B 1, Satz 5),}\\ &= B\Gamma_{p+1}(B_{p+1}B_{n+1})XY && \text{(Reg. } B\text{),}\\ &=(C_1\cdot B\Gamma_{p+1})(B_{p+1}B_{n+1})YX && \text{(II B 4, Satz 1; Reg. } C\text{),}\\ &=\Gamma_{p+2}(B_{p+1}B_{n+1})YX. && \text{(Satz 3).}\end{aligned}$$

SATZ 7. *Wenn XY Etwase sind, und Y das Kommutativgesetz*

$$\vdash C_1B_{m+1}Y=BY\cdot B_n$$

erfüllt; dann

$$\vdash \Gamma_{p+1}(B_pB_{m+1}\cdot X)Y=B_{p+1}Y\cdot B_pB_n\cdot X.$$

Beweis: Nach den Voraussetzungen,

$$\begin{aligned}\vdash \Gamma_{p+1}(B_pB_{m+1}\cdot X)Y &= \Gamma_{p+2}(B_{p+1}B_{m+1})YX && \text{(Satz 6),}\\ &= B_{p+1}(C_1B_{m+1}Y)X && \text{(Satz 5),}\\ &= B_p(C_1B_{m+1}Y)\cdot X && \text{(II B 1, Satz 5; II B 4, Def. 1),}\\ &= B_p(BY\cdot B_n)\cdot X && \text{(Hp.),}\\ &= B_{p+1}Y\cdot B_pB_n\cdot X && \text{(II B 4, Satz 6; II B 1, Satz 5).}\end{aligned}$$

§ 3. *Darstellung der allgemeinen Kombinationen.*

Festsetzung 1. Ein Ausdruck X der Form

$$(\mathfrak{R}IY_1Y_2\cdots Y_px_1x_2\cdots x_n),$$

wo die Y_i Etwase sind, *reduziert sich formal* auf einen Ausdruck Z, wenn mit Behandlung der Y_i als Variablen eine Reduktion von X auf Z sich durchführen lässt; oder, falls man es genauer haben will, wenn der Ausdruck $(\mathfrak{R}Ix_1x_2\cdots x_{n+p})$ sich auf ein solches Z' reduziert, dass durch Einsetzung von Y_i statt x_i für $i=1,2,\cdots,p$, und von x_{i-p} statt x_i für $i=p+1$, $p+2,\cdots,p+n$ in Z', der Ausdruck Z erzielt wird.

SATZ 1. *Ist X ein Kombinator, so gibt es ein S der Form ($\mathfrak{R}IBCWK$), das sich auf X formal reduziert.*

Beweis: Ersetzen wir in dem gegebenen Kombinator B, C, W, K durch x_1, x_2, x_3 bzw. x_4, so erzeugen wir eine Kombination Z von x_1, x_2, x_3, x_4. Nach § 1, Sätzen 6 und 7 gibt es ein $\mathfrak{R}$, sodass ($\mathfrak{R}Ix_1x_2x_3x_4$) sich auf Z reduziert. Daher reduziert ($\mathfrak{R}IBCWK$) sich formal auf X, w. z. b. w.

SATZ 2. *X sei eine Kombination von Kombinatoren und Variablen $x_1, x_2, \cdots, x_n$, sodass*

a) *X auf einen ähnlichen X' durch einen einzigen Reduktionsprozess reduziert wird,*

b) *es ein S gibt, näm.*

(A) $$S \equiv \mathfrak{R}IY_1Y_2 \cdots Y_p,$$

wo jedes Y_i entweder B, C, W oder K ist, sodass der Ausdruck ($Sx_1x_2 \cdots x_n$) sich auf X formal reduziert.

Dann gibt es ein S', nämlich

(B) $$S' \equiv \mathfrak{R}'IY_1'Y_2' \cdots Y_q'$$

wo jedes Y_i' entweder B, C, W oder K ist, sodass

α) $\vdash S' = S,$

β) *der Ausdruck ($S'x_1x_2 \cdots x_n$) sich auf X' formal reduziert.*

Beweis: $Y_0', Y_1', Y_2', \cdots, Y_q'$ seien die sämtlichen in X vorkommenden Grundkombinatoren (B, C, W oder K), und zwar so, dass jeder der Kombinatoren B, C, W, K unter diesen Y_i' genau so oft erscheint, wie in X selbst. Die Anordnung dieser Kombinatoren unter den Y_i' bleibt für jetzt gleichgültig. Die Y_i' kommen natürlich—abgesehen von ihrer Häufigkeit—unter den $Y_1, Y_2, \cdots, Y_p$ vor.

Behandeln wir nunmehr die $Y_1, Y_2, \cdots, Y_p$ formal als Variable, so schliessen wir die Folgenden:

1) Die Folge $IY_0'Y_1' \cdots Y_q'x_1x_2x_3 \cdots$ ist eine Umwandlung der Folge $IY_1Y_2 \cdots Y_px_1x_2x_3 \cdots,$

2) Wenn wir die durch X bestimmte Folge wie folgt schreiben, $y_0y_1y_2y_3 \ldots$, wo y_0 entweder ein Y_i oder eine Variable ist, und y_i, $i > 0$, eine Kombination von $Y_1, Y_2, \cdots Y_p$ und Variablen ist, so ist die Folge

(1) $$Iy_0y_1y_2 \cdots$$

das Produkt der eben erwähnten Umwandlung und eine Folge derselben Form wie (1).

Nun bezeichne ich mit Ω bzw. $\mathfrak{R}_1$ zwei normale Kombinatoren, sodass Ω bzw. $\mathfrak{R}_1 I$ dieser Umwandlung bzw. der zuletzt erwähnten Folge entsprechen. Dann bemerken wir: 1) $(\Omega \cdot \mathfrak{R}_1)$ entspricht der Folge (1) (II C 2, Satz 2); 2) wir dürfen annehmen dass $\mathfrak{R}_1$ und also $(\Omega \cdot \mathfrak{R}_1)$* kein Glied der Form B_n enthält (weil $\mathfrak{R}_1$ normal ist und $(\mathfrak{R}_1 I)$ einer Folge der Form (1) entspricht—vgl. Beweis von § 1, Satz 6); 3) wir dürfen ferner annehmen, dass $\mathfrak{R}$ kein Glied der Form B_n hat (denn wenn $\vdash \mathfrak{R} = \mathfrak{R}^* \cdot B_n$, $\mathfrak{R}^*$ normal, so können wir in den Satz $\mathfrak{R}$ durch $\mathfrak{R}^*$ ersetzen). Daraus folgt

$$
\begin{aligned}
& \vdash \mathfrak{R} = \Omega \cdot \mathfrak{R}_1 && (\S\,1, \text{ Satz } 6). \\
\therefore\quad & \vdash S = (\Omega \cdot \mathfrak{R}_1) I Y_1 Y_2 \cdots Y_p && \\
& \quad = \mathfrak{R}_1 I Y_0' Y_1' \cdots Y_q' && (\text{nach der Bedeutung von } \Omega).
\end{aligned}
$$

Weiterhin reduziert der Ausdruck

$$(\mathfrak{R}_1 I Y_0' Y_1' \cdots Y_q' x_1 x_2 \cdots x_n) \tag{2}$$

sich formal auf X. (Nach der Bedeutung von $\mathfrak{R}_1$, § 1, Satz 7).*

Wir unterscheiden nun zwei Fälle; näm.—

I. Die Reduktion von X auf X' volzieht sich in dem ersten Sinne.

II. Die Reduktion von X auf X' vollzieht sich in dem zweiten Sinne.

Fall I. Hier sei Y_0' der erste in X vorkommende Grundkombinator. Dann erscheint Y_0' in X nur an der ersten Stelle. Deshalb ist X eine normale Kombination von $Y_0', Y_1', Y_2', \cdots, Y_q'$ und Variablen. Es gibt also, nach § 1, Satz 5, ein $\mathfrak{R}_2$ und ein $\mathfrak{B}$, sodass

$$
\begin{aligned}
& \vdash \mathfrak{R}_1 = B\mathfrak{R}_2 \cdot \mathfrak{B}. && \\
\therefore\quad & \vdash \mathfrak{R}_1 I Y_0' = (\mathfrak{R}_2 \cdot \mathfrak{B} I) Y_0' && \\
(3)\quad & \quad = (\mathfrak{R}_2 \cdot \mathfrak{B} I \cdot B Y_0') I && (\S\,1, \text{ Satz } 3).
\end{aligned}
$$

Nun betrachten wir Y_0' wieder als Kombinator und definieren:

a) $\mathfrak{R}' \equiv$ eine normale Form von $(\mathfrak{R}_2 \cdot \mathfrak{B} I \cdot B Y_0')$† ohne Glieder dor Form B_n,

b) $$S' \equiv \mathfrak{R}' I Y_1' Y_2' \cdots Y_q',$$

so folgt $\vdash S' = S$. Also ist die Bedingung α) erfüllt.

Dieses $\mathfrak{R}' I$ entspricht, wenn wir $Y_1', Y_2', \cdots, Y_q'$ formal betrachten, der durch X' bestimmten Folge. Denn ich habe gezeigt, dass der Ausdruck (2)

---

* **Sogar wenn es auf die Normalform gebracht wird.**

† **Dies ist regulär nach § 1, Sätzen 1 und 2.**

sich formal auf X reduziert. In dieser Reduktion betrachten wir Y_0' nunmehr nicht als Variable, sondern als Kombinator; dabei wird nichts in der Reduktion geändert. Die Reduktion lässt sich doch eine Stufe weiter auf X' durchführen (nach Hp. a). Aber weil

$$\vdash \mathfrak{N}' = \mathfrak{N}_1 I Y_0' \qquad \text{(aus (3))},$$

und die beiden Seiten dieser Gleichung Folgen lauter Variablen entsprechen, so entsprechen sie derselben Folge (II C 1, Satz 11).

Dass die Bedingung β) erfüllt ist, folgt daraus nach § 1, Satz 7.

Fall II. Hier soll Y_0' den Kombinator bezeichnen, welcher durch die Reduktion von X auf X' eliminiert wird. Er nehme in X die $(r+1)$te Stelle ein, wo $r > 0$ nach der Voraussetzung dieses Falles ist.

Nach II D 6, Satz 1, gibt es $\mathfrak{K}_1$, $\mathfrak{W}_1$, $\mathfrak{C}_1$, und $\mathfrak{B}_1$ derart, dass

$$\vdash \mathfrak{N}_1 = \mathfrak{K}_1 \cdot \mathfrak{W}_1 \cdot \mathfrak{C}_1 \cdot \mathfrak{B}_1. \tag{4}$$

Aber nach der Voraussetzung über $Y_0', Y_1', \cdots, Y_q'$ kann $\mathfrak{K}_1$ kein K_i für $i \leqq q+1$ und $\mathfrak{W}_1$ kein W_j für $j \leqq q+1$ enthalten, also wird

$$\vdash \mathfrak{K}_1 \cdot \mathfrak{W}_1 = B_{q+1}(\mathfrak{K}_2 \cdot \mathfrak{W}_2). \tag{5}$$

Auch entspricht $\mathfrak{C}_1$ einer Permutationsfolge, welche in zwei Faktoren zerlegt werden kann, wie folgt: der erste Faktor lässt Y_0' invariant, aber ordnet $Y_1', Y_2', \cdots, Y_q'$ und die Variablen in die Anordnung, die sie in X haben, an; der zweite Faktor setzt Y_0' an die Stelle, die es in X hatt, aber lässt die Anordnung von $Y_1', \cdots, Y_q'$ und die Variablen unter sich selbst, unverändert bleiben. Dem ersten Faktor entspricht ein $\mathfrak{C}$, dessen Glieder alle C_i mit $i > 1$ sind, also ein $\mathfrak{C}$ von der Form $B\mathfrak{C}_2$; dem zweiten Faktor entspricht Γ_r $(r > 0)$. Also (II D 5, Satz 7).

$$\vdash \mathfrak{C}_1 = B\mathfrak{C}_2 \cdot \Gamma_r. \tag{6}$$

Daher (aus (4) (5) (6))

$$\vdash \mathfrak{N}_1 = B(B_q\mathfrak{K}_2 \cdot B_q\mathfrak{W}_2 \cdot \mathfrak{C}_2) \cdot \Gamma_r \cdot \mathfrak{B}. \tag{7}$$

Nun erscheint Y_0' nach Hp. (a) und Definition am Anfang eines in X eingeklammerten Teilausdrucks; die Anzahl der Glieder ausser Y_0' dieses Teilausdrucks sei $m+1$. Dann (vgl. den Beweis von II C 3, Satz 3, und II D 3, Satz 1) gibt es $\mathfrak{B}_2$ und $\mathfrak{B}_3$ derart, dass

$$\vdash \mathfrak{B} = B_{r+1}\mathfrak{B}_2 \cdot B_r B_{m+1} \cdot \mathfrak{B}_3. \tag{8}$$

Weil die Glieder von Γ_r alle $C_1, C_2, \cdots$ oder C_r sind, so kann $B_{r+1}\mathfrak{B}_2$ mit allen diesen Gliedern, also mit Γ_r selbst, vertauscht werden (II D 2, Satz 5a).

Daher $\vdash \mathfrak{R}_1 = B(B_q\mathfrak{K}_2 \cdot B_q\mathfrak{W} \cdot \mathfrak{C}_2) \cdot \Gamma_r \cdot B_{r+1}\mathfrak{B}_2 \cdot B_rB_{m+1} \cdot \mathfrak{B}_3,$

$$= B(B_q\mathfrak{K}_2 \cdot B_q\mathfrak{W}_2 \cdot \mathfrak{C}_2 \cdot B_r\mathfrak{B}_2) \cdot \Gamma_r \cdot B_rB_{m+1} \cdot \mathfrak{B}_3$$

(9) $$= B\mathfrak{R}_2 \cdot \Gamma_r \cdot B_rB_{m+1} \cdot \mathfrak{B}_3.$$

wenn ich nur definiere:

$$\mathfrak{R}_2 \equiv B_q\mathfrak{K}_2 \cdot B_q\mathfrak{W}_2 \cdot \mathfrak{C}_2 \cdot B_r\mathfrak{B}_2.$$

Daher $\vdash \mathfrak{R}_1IY_0' = (B\mathfrak{R}_2 \cdot \Gamma_r \cdot B_rB_{m+1} \cdot \mathfrak{B}_3)IY_0'$

$$= B\mathfrak{R}_2(\Gamma_r((B_rB_{m+1} \cdot \mathfrak{B}_3)I))Y_0'$$

(10) $$= \mathfrak{R}_2(\Gamma_r(B_{r-1}B_{m+1} \cdot \mathfrak{B}_3I)Y_0').$$

Weil nach Hp. a) und Definition von Y_0' eine Reduktion durch Y_0' wirklich stattfindet, so muss $m \geqq 2$ sein, wenn $Y_0'B$ oder C ist, und $m \geqq 1$, wenn $Y_0'W$ oder K ist. Infolgedessen muss es nach II D 2, Satz 2, und den kommutativen Axiomen ein n geben, wofür $\vdash C_1B_{m+1}Y_0' = BY_0' \cdot B_n$. Also

$$\vdash \Gamma_r(B_{r-1}B_{m+1} \cdot \mathfrak{B}_3I)Y_0' = B_rY_0' \cdot B_{r-1}B_n \cdot (\mathfrak{B}_3I) \quad (\S\,2,\ \text{Satz } 7);$$

also, wenn wir dies in (10) einsetzen,

$$\vdash \mathfrak{R}_1IY_0' = \mathfrak{R}_2(B_rY_0' \cdot B_{r-1}B_n \cdot \mathfrak{B}_3I)$$

(11) $$= (\mathfrak{R}_2 \cdot B_{r+1}Y_0' \cdot B_rB_n \cdot \mathfrak{B}_3)I \quad (\S\,1,\ \text{Satz } 3).$$

Definiere ich nun

a) $$\mathfrak{R}' \equiv \mathfrak{R}_2 \cdot B_{r+1}Y_0' \cdot B_rB_n \cdot \mathfrak{B}_3,$$

b) $$S' \equiv \mathfrak{R}'IY_1'Y_2' \cdots Y_q',$$

so folgt aus (11) und (A), dass $\vdash S = S'$.

Dass die Bedingung β) erfüllt ist, folgt hier genau wie im Fall I.

SATZ 3. *Ist X ein solcher Kombinator, dass*

(1) $$(Xx_1x_2 \cdots x_n)$$

sich auf eine Kombination von $x_1, x_2, \cdots, x_n$ reduziert; dann lässt X sich in eine $(\mathfrak{R}I)$ umformen und zwar so, dass $(\mathfrak{R}Ix_1x_2 \cdots x_n)$ sich auf die gegebene Kombination reduziert.

Beweis: Nach den Voraussetzungen gibt es eine Reihe von Ausdrücken $X_1, X_2, \cdots, X_m$ derart, dass 1) X_{i+1} sich aus X_i durch einen einzigen Reduktionsprozess erzielt, 2) X_1 mit dem Ausdruck (1) identisch ist, 3) X_m eine Kombination von $x_1, x_2, \cdots, x_n$ ist.

Wir können nun diesen X_i eine Reihe von Kombinatoren $S_1, S_2, \cdots, S_m$ zuordnen und zwar so dass

a) Jedes S_i in der Form (A) (s. Satz 2) steht,
b) $(S_ix_1x_2 \cdots x_n)$ sich auf X_i formal reduziert,
c) $\vdash S_{i+1} = S_i$.

In der Tat gilt als S_1 der in Satz 1 ausgestellte Kombinator; und aus Satz 2 folgt, dass aus einem gegebenen S_i, $(i < m)$ ein S_{i+1} konstruiert werden kann.

In dieser Weise haben wir ein S_m, etwa

$$(2) \qquad S_m \equiv \mathfrak{N}_m I Y_1 Y_2 \cdots Y_p \quad (Y_i \equiv B, C, W \text{ oder } K, \mathfrak{N}_m \text{ normal})$$

sodass $(S_m x_1 x_2 \cdots x_n)$ sich *formal* auf eine Kombination lauter Variablen reduziert. In dieser Reduktion müssen freilich alle $Y_1, Y_2, \cdots, Y_p$ ausfallen. Also wenn $\mathfrak{N}_m$ auf die normale Form gebracht wird, gilt

$$\vdash \mathfrak{N}_m = K_p \cdot K_{p-1} \cdots K_1 \cdot \mathfrak{N}_m'.$$

Infolgedessen	$\vdash S_m = \mathfrak{N}_m' I$	(aus (2), II B 3).
Aber	$\vdash S_m = S_1$	(aus c)),
	$= X$	(Bedeutung von S_1).
$\therefore$	$\vdash X = \mathfrak{N}_m' I.$	w. z. b. w.

SATZ 4. *Wenn zwei Kombinatoren Y_1 und Y_2 derselben Folge lauter Variablen entsprechen;*

dann $\qquad \vdash Y_1 = Y_2.$

Beweis: Nach Satz 3 gibt es $\mathfrak{N}_1$ und $\mathfrak{N}_2$, sodass

$$\vdash Y_1 = \mathfrak{N}_1 I \qquad \vdash Y_2 = \mathfrak{N}_2 I,$$

und die beiden Kombinatoren $(\mathfrak{N}_1 I)$ und $(\mathfrak{N}_2 I)$ auch derselben Folge entsprechen. Wir können ohne Beschränkung der Allgemeinheit annehmen, dass $\mathfrak{N}_1$ und $\mathfrak{N}_2$ normal und ohne Glieder der Form B_n sind.

Dann	$\vdash \mathfrak{N}_1 = \mathfrak{N}_2$	(§ 1, Satz 6).
Also	$\vdash Y_1 = Y_2.$	w. z. b. w.

SATZ 5. *Wenn zwei eigentliche Kombinatoren Y_1 und Y_2 dieselbe Kombination von lauter Variablen darstellen, dann* $\vdash Y_1 = Y_2$.

Beweis: Klar aus Satz 4.

§ 4. *Die Substitutionsprozesse.*

Zum Schluss gebe ich hier einige Sätze über die Verhältnisse der Substitutionsprozesse zu den Kombinatoren. Die Bewiese gebe ich nur kurz, weil sie meistens nur Rechnungsübungen sind.

Die Substitutionsprozesse lassen sich zunächst durch Kombinationen von Variablen darstellen. Z. B. betrachten wir den Ausdruck:

$$(u x_1 (v x_2 x_3) x_4).$$

Wenn u und v Grundfunktionen sind, so bedeutet dies eine gewisse aus einer

Verknüpfung von u und v erzeugte Funktion von x_1, x_2, x_3, x_4. Aber wir können ihn auch,—wenn wir u und v für Variablen halten—als eine Funktion von u und v betrachten, welche für bestimmte Werte von u und v jene Funktion von x_1, x_2, x_3, x_4 darstellt—d. h. als den Verknüpfungsprozess selbst. Diese Auffassung ist naturgemäss, weil nach der Ausdeutung von Anwendung der Ausdruck für irgendeine bestimmten Werte von u, v, x_1, x_2, x_3, x_4 die mit der Auffassung verträgliche Aussage bedeutet. Der Ausdruck lässt sich ferner in

$$(C_1 \cdot BB_2)uvx_1x_2x_3x_4$$

umformen. Von unserem Gesichtspunkte aus ist also $(C_1 \cdot BB_2)$ der Substitutionsprozess selbst—eine Funktion, welche aus u und v die Funktion $((C_1 \cdot BB_2)uv)$ liefert, wo diese letzte die Funktion ist, welche aus x_1, x_2, x_3 die eben geschilderte Aussage liefert. In diesem Sinne können wir sagen, dass $(C_1 \cdot BB_2)$ den betreffenden Substitutionsprozess *darstellt.*

Von diesem Gesichtspunkt aus haben wir die folgenden Satze:

SATZ 1. *Jede Umwandlung im Sinne von Abschnitt A lässt sich durch ein Ω darstellen.*

SATZ 2. *Die Einsetzung von einer Funktion als Funktion von n Variablen an die Stelle der $(m+1)$ ten Variablen einer zweiten wird durch $(\Gamma_m \cdot B_mB_n)$ dargestellt.*

SATZ 3. *Sind die Substitutionsprozesse wie in den Sätzen 1 und 2 dargestellt, dann gestalten sie Ausdrücke der Form $(Yu_1u_2\cdots u_n)$, wo Y eine eigentliche Kombination von Ordnung nicht zu gross ist, in andere Ausdrücke derselben Form um.*

Beweis: Für eine Umwandlung gilt

$$\begin{aligned}\vdash \Omega(Yu_1u_2\cdots u_n) &= B_n\Omega Yu_1u_2\cdots u_n\\ &= (B_{n-1}\Omega \cdot Y)u_1u_2\cdots u_n.\end{aligned}$$

Für Zusammensetzungen: es sei $Z \equiv \Gamma_p \cdot B_pB_q$; dann

$$\begin{aligned}\vdash Z(Xu_1u_2\cdots u_m)(Yv_1v_2\cdots v_n)&\\ &= (B_mZ \cdot BX)Iu_1u_2\cdots u_m(Yv_1v_2\cdots v_n)\\ &= (\Gamma_m \cdot B_mB_n \cdot B_mZ \cdot BX)IYu_1u_2\cdot u_mv_1v_2\cdots v_n\\ &= Uu_1u_2\cdots u_mv_1v_2\cdots v_m,\end{aligned}$$

wo U eigentlich ist, wenn nur X und Y eigentlich sind, und q gross genug ist, sodass $Yx_1x_2\cdots x_{n+q}$ sich auf eine Kombination lauter Variablen reduziert. In der Tat reduziert $(Uu_1u_2\cdots u_mv_1v_2\cdots v_nx_1x_2\cdots x_{p+q})$ sich

11

auf $(Xu_1u_2 \cdots u_m x_1 x_2 \cdots x_p(Yv_1v_2 \cdots v_n x_{p+1} \cdot x_{p+q}))$. Die Bedingung auf Y ist erfüllt, wenn wir es mit einem Substitutionsprozess zu tun haben.

SATZ 4. *X sei eine Kombination von Variablen und gewissen Etwasen $u_1, u_2, \cdots, u_m$. Dann gibt es einen Kombinator Y, sodass*

$$\vdash Yu_1u_2u_mx_1x_2 \cdots x_n = X, \quad (u_i \text{ als Variable behandelt}).$$

Gibt es weiter einen Kombinator Z, sodass

$$\vdash Z_1v_2v \cdots v_px_1x_2 \cdots x_n = X \quad (v_i \text{ als Variable betrachtet}),$$

wo $\quad \vdash Iv_1v_2 \cdots v_p = VIu_1u_2 \cdots u_m, \quad (V \text{ ein Kombinator}),$

(oder umgekehrt),
dann $\quad \vdash Yu_1u_2 \cdots u_m = Zv_1v_2 \cdots v_p.$

*Beweis:** Wenn die $v_1v_2 \cdots v_p$ dieselbe Reihe von Etwasen bildet, wie $u_1u_2, \cdots, u_m$, so folgt der Satz aus § 3, Satz 5. Sonst

$$\vdash Zv_1v_2 \cdots v_p = (V \cdot BZ)Iu_1u_2 \cdots u_p,$$

und $\quad \vdash (V \cdot BZ)I = Y \quad$ (§ 3, Satz 5).

$\therefore \quad \vdash Zv_1v_2 \cdots v_p = Yu_1u_2 \cdots u_m \quad$ w. z. b. w.

* Der Beweis des ersten Teils des Satzes ist klar (§ 1, Satz 8).

A Test for the Type of Irrationality Represented by a Periodic Ternary Continued Fraction.

By J. B. Coleman.

1. *Introduction.* Let $(\overline{p_1, q_1; p_2, q_2; \cdots; p_k, q_k;} \cdots)$ denote a purely periodic ternary continued fraction,* of period $k \geqq 4$,† the partial quotient pairs being real numbers. Let D_1 denote the determinant

$$\begin{vmatrix} -p_1 & q_1 & 1 & 0 & 0 & \cdots & 0 & 0 & 0 & (-1)^k & 0 \\ 1 & -p_2 & q_2 & 1 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & -p_3 & q_3 & 1 & \cdots & 0 & 0 & 0 & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 1 & -p_{k-1} & q_{k-1} & 1 \\ 0 & (-1)^k & 0 & 0 & 0 & \cdots & 0 & 0 & 1 & -p_k & q_k \\ (-1)^k & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & -1 \end{vmatrix}$$

and D_2 be the determinant derived from D_1 by replacing $(-1)^k$ by $(-1)^{k+1}$ where it occurs in the first row and in the second column, and replacing -1 by 1 in the last column.

In this paper I prove that the vanishing of D_1 or of D_2 is a necessary and sufficient condition for the reducibility of the characteristic equation, when p_i and q_i are rational integers, in which case the ternary continued fraction represents a rational number or a quadratic irrationality.‡ If p_i and q_i are any real numbers, the vanishing of D_1 or of D_2 is a sufficient condition for the reducibility of the characteristic equation.

We proceed to prove the above statements by first finding determinant forms for the convergents and other expressions involved in the characteristic equation. This is done in sections 2-6. In sections 7-11, following the general proof, are some corollaries and numerical examples.

2. Let three sequences satisfying the recursion formula

$$W_n = q_n W_{n-1} + p_n W_{n-2} + W_{n-3} \tag{1}$$

* References for previous history: C. G. J. Jacobi, *Werke*, Vol. 6, p. 385; O. Perron, *Mathematische Annalen*, Vol. 64, p. 1; D. N. Lehmer, *Proceedings of the National Academy of Sciences*, Vol. 4, p. 360; H. P. Daus, *American Journal of Mathematics*, Vol. 51, p. 67; O. Perron, "Die Lehre von den Kettenbrüchen."

† If $k < 4$ the expressions for D_1 and D_2 must be interpreted as shown in section 9.

‡ For convergence conditions see O. Perron, article cited.

with the initial values

(1a) $$(0,0,1) \quad (0,1,0) \quad (1,0,0)$$

be denoted, respectively, by C_n, B_n, A_n.

The characteristic equation is

(2) $$\rho^3 - M\rho^2 + N\rho - 1 = 0,$$

in which

(3a) $$M = A_{k-2} + B_{k-1} + C_k,$$

(3b) $$N = (A_{k-2}, B_{k-1}) + (A_{k-2}, C_k) + (B_{k-1}, C_k).$$

$(B_{k-1}, C_k) = B_{k-1}C_k - B_kC_{k-1}$, &c.

Since by a theorem of Lehmer's * the roots, $\sigma_{1,1}$, and $\sigma_{2,1}$, of the cubic equation representing the expansion are related to ρ_1, the principal root of (2), as follows;

(4a) $$\sigma_{1,1} = (B_k\rho_1 + A_kB_{k-2} - A_{k-2}B_k)/(A_k\rho_1 + A_{k-1}B_k - A_kB_{k-1})$$

(4b) $$\sigma_{2,1} = (C_{k-1}\rho_1 + C_{k-2}A_{k-1} - C_{k-1}A_{k-2})/(A_{k-1}\rho_1 + A_kC_{k-1} - A_{k-1}C_k),$$

the rationality or type of irrationality represented by the continued fraction may be determined from a discussion of the characteristic equation. Since (2) is of the third degree, if it is reducible when M and N are rational integers, it must have a factor $\rho - 1$ or $\rho + 1$. Hence the necessary and sufficient conditions for reducibility under these conditions, are

(5a) $$-M + N = 0 \quad \text{or}$$

(5b) $$M + N + 2 = 0.$$

If M and N are any real numbers, conditions (5a) or (5b) will be sufficient to insure reducibility.

3. To find determinant forms for A_n, B_n and C_n.

Let $$M\begin{pmatrix} a_1, a_2, a_3, \cdots, a_n \\ b_2, b_3, \cdots, b_n \end{pmatrix}_n$$

denote the determinant

(6) $$\begin{vmatrix} a_1 & 1 & 0 & 0 & 0 & \cdot & \cdot & \cdot & \cdot \\ -b_2 & a_2 & 1 & 0 & 0 & \cdot & \cdot & \cdot & \cdot \\ 1 & -b_3 & a_3 & 1 & 0 & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & 0 & 1 & -b_{n-1} & a_{n-1} & 1 \\ \cdot & \cdot & \cdot & \cdot & 0 & 0 & 1 & -b_n & a_n \end{vmatrix}.$$

* D. N. Lehmer, *loc. cit.*

Following the usage in ordinary continued fractions we shall call (6) a *continuant.*

Expanding (6) according to the elements of its last row,

$$(7a)\qquad M\begin{pmatrix} a_1, a_2, a_3, \cdots, a_n \\ b_2, b_3, b_4, \cdots, b_n \end{pmatrix}_n = a_n M\begin{pmatrix} a_1, a_2, \cdots, a_{n-1} \\ b_2, b_3, \cdots, b_{n-1} \end{pmatrix}_{n-1} + b_n M\begin{pmatrix} a_1, a_2, \cdots, a_{n-2} \\ b_2, \cdots, b_{n-2} \end{pmatrix}_{n-2}$$
$$+ M\begin{pmatrix} a_1, a_2, \cdots, a_{n-3} \\ b_2, b_3, \cdots, b_{n-3} \end{pmatrix}_{n-3}, \qquad (n > 3).$$

By (1) and (1a) we have directly

$$(7b)\qquad A_2 = M(q_2), \quad A_3 = M\begin{pmatrix} q_2, q_3 \\ p_3 \end{pmatrix}_2, \quad A_4 = M\begin{pmatrix} q_2, q_3, q_4 \\ p_3, p_4 \end{pmatrix}_3.$$

From (7a) and (1) we have immediately

$$(8)\qquad A_n = M\begin{pmatrix} q_2, q_3, q_4, \cdots, q_n \\ p_3, p_4, \cdots, p_n \end{pmatrix}_{n-1}, \qquad (n = 2, 3, \cdots, n),$$

for any n, provided it is true for three successively lower values of n. But by (7b) the relation (8) is true for $n = 2, 3, 4$; hence, by induction, it is true for all values of n.

In the same way we find *

$$(9)\qquad B_n = p_1 A_n + M\begin{pmatrix} q_3, q_4, \cdots, q_n \\ p_4, p_5, \cdots, p_n \end{pmatrix}_{n-2}, \qquad (n = 3, 4, \cdots).$$

$$(10)\qquad C_n = M\begin{pmatrix} q_1, q_2, q_3, \cdots, q_n \\ p_2, p_3, \cdots, p_n \end{pmatrix}_n, \qquad (n = 1, 2, 3, \cdots).$$

4. By the recursion formulae (1) and (1a) it is found that

$$(A_{-2}, B_{-1}) = 1, \quad (A_{-1}, B_0) = 0, \quad (A_0, B_1) = 0, \quad (A_1, B_2) = 1,$$

and $$(A_n, B_{n+1}) = -p_{n+1}(A_{n-1}, B_n) - q_n(A_{n-2}, B_{n-1}) + (A_{n-3}, B_{n-2}), \qquad (n = 1, 2, 3, \cdots).$$

By direct calculation it is found that

$$(A_2, B_3) = M(-p_3)_1, \quad (A_3, B_4) = M\begin{pmatrix} -p_3, -p_4 \\ -q_3 \end{pmatrix}_2,$$
$$(A_4, B_5) = M\begin{pmatrix} -p_3, -p_4, -p_5 \\ -q_3, -q_4 \end{pmatrix}_3.$$

Hence

$$(11)\qquad (A_n, B_{n+1}) = M\begin{pmatrix} -p_3, -p_4, \cdots, -p_{n+1} \\ -q_3, -q_4, \cdots, -q_n \end{pmatrix}_{n-1}, \qquad (n = 2, 3, \cdots),$$

by inductive reasoning similar to that employed in deriving (8).

* A set of continuants similar to (8), (9) and (10) may be written for the convergents involved in quaternary, or in n-ary, continued fractions. In every case the proof, by induction, involves assuming that the convergent is represented by its continuant for n successive orders and proving it true for the next higher order.

By (1) and by definition

$$(A_n, C_{n+2}) = q_{n+2}(A_n, C_{n+1}) - (A_{n-1}, C_n).$$

Hence

$$(A_n, C_{n+2}) = q_{n+2} M \begin{pmatrix} -p_2, -p_3, \cdots, -p_{n+1} \\ -q_2, \cdots, -q_n \end{pmatrix}_n - M \begin{pmatrix} -p_2, -p_3, \cdots, -p_n \\ -q_2, \cdots, -q_{n-1} \end{pmatrix}_n$$

since by a process similar to that used in deriving (11) we have

$$(12) \qquad (A_n, C_{n+1}) = M \begin{pmatrix} -p_2, -p_3, -p_4, \cdots, -p_{n+1} \\ -q_2, -q_3, \cdots, -q_n \end{pmatrix}_n, \quad (n = 1, 2, \cdots).$$

Also from similar considerations

$$(13) \qquad (B_n, C_{n+1}) = M \begin{pmatrix} -p_1, -p_2, \cdots, -p_{n+1} \\ -q_1, -q_2, \cdots, -q_n \end{pmatrix}_{n+1}, \qquad (n = 0, 1, 2, \cdots).$$

5. Proof that $D_2 = M + N + 2$.

Expanding D_2 in terms of the elements in the last row we obtain two determinants of order k. We shall designate by D_3 the minor of 1, and by D_4 the minor of $(-1)^k$, in the last row. Next we expand D_3 in terms of the elements of the last row and last column, by Cauchy's method. Applying (13) three times, (8) twice, and (9) three times to terms of this expansion, we obtain

$$(14) \qquad \begin{aligned} D_3 = &- p_k(B_{k-2}, C_{k-1}) - q_{k-1}(B_{k-3}, C_{k-2}) + (B_{k-4}, C_{k-3}) \\ &+ q_{k-1}B_{k-2} + p_{k-1}B_{k-3} + B_{k-4} + (A_{k-2}, B_{k-1}) + 1. \end{aligned}$$

Now by the recursion formulae (1) for B_n and C_n and by definition, (B_{k-1}, C_k) reduces to $-p_k(B_{k-2}, C_{k-1}) - (B_{k-3}, C_{k-1})$ and also (B_{k-3}, C_{k-1}) reduces to $q_{k-1}(B_{k-3}, C_{k-2}) - (B_{k-4}, C_{k-3})$. Hence the first three terms of (14) reduce to (B_{k-1}, C_k). By (I) for B_n the next three terms of (14) reduce to B_{k-1}. Substituting these terms in (14) it becomes

$$(15) \qquad D_3 = (B_{k-1}, C_k) + B_{k-1} + (A_{k-2}, B_{k-1}) + 1.$$

Next expand D_4 in terms of the elements of the last row and next to the last column, by Cauchy's method. Applying (12) twice, (10) five times and (8) once to the expansion gives

$$(16) \qquad \begin{aligned} D_4 = q_k(A_{k-2}, C_{k-1}) &- (A_{k-3}, C_{k-2}) + p_k C_{k-2} + q_{k-1}q_k C_{k-2} \\ &+ q_k p_{k-1} C_{k-3} + q_k C_{k-4} + C_{k-3} + A_{k-2} + 1. \end{aligned}$$

By recursion formulae (1) for A_n and C_n the first two terms of (16) reduce to (A_{k-2}, C_k). By (1) for C_k and C_{k-1} the next five terms become C_k. Substituting these values in (16) it becomes

(17) $$D_4 = (A_{k-2}, C_k) + C_k + A_{k-2} + 1.$$

Combining (15) and (17) gives

$$D_2 = D_3 + D_4 = C_k + B_{k-1} + A_{k-2} + (B_{k-1}, C_k) + (A_{k-2}, C_k) + (A_{k-2}, B_{k-1}) + 2.$$

Hence by (3) and (4),

$$D_2 = M + N + 2.$$

6. To show that

(18) $$D_1 = C_k + B_{k-1} + A_{k-2} - (B_{k-1}, C_k) - (A_{k-2}, C_k) - (A_{k-2}, B_{k-1})$$

it is not necessary to expand it completely as was done for D_2. The elements of the two determinants correspond except that three elements of each are replaced in the other by the same elements with their signs changed. Hence the expansion of D_1 may be obtained from that of D_2 by making appropriate changes of sign. The result of making these changes of sign in the preceding section produces (18). Thus by (3) and (4), $D_1 = M - N$.

From (5) and (6) it is now evident that the proof of the original statements is complete, i. e., that for the reducibility of the characteristic equation of a periodic ternary (continued) fraction, the vanishing of D_1 or D_2 is, (a) a necessary and sufficient condition where p_i and q_i are rational integers, (b) a sufficient condition when p_i and q_i are any real numbers.

7. Since D_1 and D_2 are linear in any particular p_i and q_i, it is obvious that if $k-1$ pairs of partial quotients be selected arbitrarily, the remaining pair may be selected in an infinite number of ways so as to make the characteristic equation reducible, either by the root 1 or -1.

In the same way D_1 and D_2 are linear in any q_i and p_{i+1}, so that the same statement may be made for such a pair as for the pair p_i and q_i.

8. Below are listed six general conditions under which the characteristic equation will be reducible. The classification is made according to the method of derivation from D_1 or D_2.

A. $p_i = -q_i$.

B. $p_i = q_i + 2$.

C. $p_1 = q_k = 0$, $p_i = -q_{i-1}$, $(i = 2, 3, \cdots, k)$.

D. $p_1 = 0$, $q_k = 2$, $p_2 + q_1 + 2 = 0$, $p_k + q_{k-1} = 2$, $p_i = -q_{i-1}$, $(i = 3, \cdots, k-1)$.

E_1. k being odd, $p_1 = q_k = 0$, $p_2 = q_1$, $p_k = q_{k-1}$, $p_i = q_{i-1} + 2$, $(i = 3, \cdots, k-1)$.

E_2. k being even, $p_1 = 2$, $q_k = 0$, $p_i = q_{i-1} + 2$, $(i = 2, 3, \cdots, k)$.

F_1. k odd, $p_1 = 0$, $q_k = -2$, $p_i = q_{i-1} + 2$, $(i = 2, 3, \cdots, k)$.
F_2. k even, $p_1 = 2$, $q_k = -2$, $p_2 = q_1$, $p_k = q_{k-1}$, $p_i = q_{i-1} + 2$,
$(i = 3, \cdots, k-1)$.

Two special cases arise amongst these. For $k = 2$, condition D becomes $p_1 = 0$, $q_2 = 2$, $q_1 = -p_2$. Also for $k = 2$, F_2 becomes $p_1 = 2$, $q_2 = -2$, $p_2 = q_1 - 2$.

Reducibility under A results from the fact that under this condition the sum of the odd columns in D_1 is equal to the sum of the even numbered columns.

Reducibility under condition B may be shown from the fact that the sum of the elements in the i-th row of D_1 is $-p_i + q_i + 2$ when k is even, and the same is true for the i-th row of D_2 when k is odd. Hence under this condition the root of the characteristic equation will be $(-1)^k$. This condition was found and proved by Lehmer.

Conditions C, D, E, F result from a consideration of the sums and differences of the two sets of alternate rows of D_1 and D_2.

9. The vanishing of D_1 or D_2, in the special cases where $k = 1, 2, 3$, may be obtained from the general formulae by observing the following. It is necessary to have the elements involving powers of -1 always occupy the three positions indicated below;

Row	Column
1	k
k	2
$k+1$	1

In case $k = 1, 2, 3$ these elements are to be added to any other elements which may occupy the same position in the determinant.

10. The character of the roots of a reducible characteristic equation when M and N are rational integers.

A. For $D_1 = 0$, or $M - N = 0$, one root of the characteristic equation is 1. The other two roots will also be rational only in case $M = N = 3$ or $M = N = -1$. When $-1 < M = N < 3$ the other roots will be imaginary. When $M = N < -1$ or $M = N > 3$, the two remaining roots will be quadratic irrationalities.

B. For $D_2 = M + N + 2 = 0$, one root of the characteristic equation is -1 and the other two are always quadratic irrationalities.

11. Numerical examples in which the characteristic equation is reducible.

A. An example in which the partial quotients are positive integers and for which D_1 vanishes.

Given $(\overline{2, 2;\ 4, 1;\ 8, 1;\ 8, 5;} \cdots)$, k, the number of pairs per period being 4.

The characteristic equation is $\rho^3 - 185\rho^2 + 185\rho - 1 = 0$. The principal root, $\rho_1 = 92 + \sqrt{8463}$.

By (4a), $$\sigma_{1,1} = \frac{11163 + 121\sqrt{8463}}{4905 + 54\sqrt{8463}}.$$

and by (4b), $$\sigma_{2,1} = \frac{2147 + 23\sqrt{8463}}{285 + 9\sqrt{8463}}.$$

B. An example containing positive and negative partial quotients, D_1 vanishing for the set.

Given $(\overline{2, 3;\ 3, -1;\ 2, 4;} \cdots)$, k being 3.

The characteristic equation is $\rho^3 - 7\rho^2 + 7\rho - 1 = 0$.

The principal root is $\rho_1 = 3 + 2\sqrt{2}$.

By (4a), $\sigma_{1,1} = (2 - \sqrt{2})/2$,

and by (4b), $\sigma_{2,1} = -(6 + 3\sqrt{2})/4$.

C. An example of the same type as *B.*

Given $(\overline{2, 3;\ 1, -1;} \cdots)$, k being 2.

The characteristic equation is $\rho^3 - 1 = 0$. The general conditions for convergence are not satisfied, so that the expansion does not give a limit for B_n/A_n or C_n/A_n.

D. An example involving positive and negative partial quotients, D_1 and D_2 both vanishing for the set.

Given $(\overline{-1, 1;\ 3, -3;} \cdots)$, k being two.

The characteristic equation is $\rho^3 - \rho^2 + \rho - 1 = 0$.

The principal root $\rho_1 = -1$.

By (4a), $\sigma_{1,1} = -1$,

and by (4b), $\sigma_{2,1} = 0$.

E. An example involving fractional partial quotients, D_1 vanishing for the set.

Given $(\overline{1/2, 2/3;\ 1/3, -4/5;} \cdots)$, k being 2.

The characteristic equation is $\rho^3 - 3/10\rho^2 + 3/10\rho - 1 = 0$.

This equation is reducible but again the conditions for convergence are not satisfied.

F. An example involving irrational partial quotients, D_1 vanishing for the set.*

Given $(\overline{\sqrt{2}, \sqrt{3};\ \sqrt{6}, \sqrt{2};\ -\sqrt{3}, \sqrt{6}};\ \cdots)$, k being 3.

The characteristic equation is $\rho^3 - 14\rho^2 + 14\rho - 1 = 0$.

The principal root is $\rho_1 = (13 + \sqrt{165})/2$.

By (4a), $\sigma_{1,1} = \sqrt{2}(\ 5 + \sqrt{165})/10$,

and by (4b), $\sigma_{2,1} = \sqrt{3}(15 + \sqrt{165})/10$.

* Under the given conditions the characteristic equation is reducible, even when irrational partial quotients are involved. However, for such a set of partial quotients, the type of irrationality defined by the continued fraction will not in general be quadratic, as was the case in example *F*.

On the Separation Property of the Roots of the Secular Equation.

By E. T. Browne.

1. *Introduction.* Let A be any square matrix, real or complex, of order n. If I is the unit matrix, $A - \lambda I$ is called the *characteristic matrix* of A; the determinant of the characteristic matrix is called the *characteristic determinant* of A; the equation obtained by equating this determinant to zero is called the *characteristic equation* of A; and the roots of this equation are called the *characteristic roots* of A. In particular, if A is real and symmetric, i. e., $a_{ij} = a_{ji}$, the characteristic equation is of great importance and is called the *secular equation* since it was first used by Laplace in the determination of the secular inequalities of the planets.

The secular equation has been widely studied and many beautiful properties of it have been discovered. For example, let us following Weber * denote by $L_i(\lambda)$ the determinant of order i standing in the upper left hand corner of $A - \lambda I$. Weber gives a proof that the roots of $L_i(\lambda) = 0$ are all real and are *separated* by the roots of $L_{i-1}(\lambda) = 0$. However, it may happen that a root ρ of multiplicity m of the latter equation is a root of multiplicity $m - 1$, m or even $m + 1$ of the former, so that if $L_{i-1}(\lambda) = 0$ has a multiple root the sense in which the previously mentioned "separation" takes place is not exactly clear. It is the purpose of this paper to study this separation property. In doing so we shall employ merely the simplest properties of algebraic equations together with a well known theorem which in the study of the characteristic equation of a matrix is one of the most useful with which the author is acquainted; viz.,

If A is a Hermitian (real symmetric) matrix of order n, there exists a unitary (real orthogonal) matrix R such that $\overline{R}'AR = N$ $(R'AR = N)$, where N has as elements in its main diagonal the (real) characteristic roots of A and zeros elsewhere.†

This theorem was used by Bromwich ‡ in his proof that if $\alpha + i\beta$ is a characteristic root of a matrix A whose elements are real or complex, and if $\rho_1 \leqq \cdot \cdot \cdot \leqq \rho_n$ are the characteristic roots of $(A + \bar{A}')/2$ and $i\mu_1, \cdot \cdot \cdot, i\mu_n$ are

* Weber, *Lehrbuch der Algebra*, Braunschweig (1898), Vol. I, pp. 307-311.

† Dickson, *Modern Algebraic Theories*, Chicago (1926), pp. 74-76; Kowalewski, *Determinantentheorie*, Berlin (1925), pp. 194-198.

‡ Bromwich, "On the Roots of the Characteristic Equation of a Linear Substitution," *Acta Mathematica*, Vol. 30 (1906), pp. 295-304.

the characteristic roots of $(A - \bar{A}')/2$, then $\rho_1 \leqq \alpha \leqq \rho_n$ and $|\beta|$ does not exceed the greatest of $|\mu_1|, \cdots, |\mu_n|$. The same theorem was employed by the author * in the proof that if λ is a charateristic root of a matrix A and if G and s are respectively the largest and smallest characteristic roots of $A\bar{A}'$, then $s \leqq \lambda\bar{\lambda} \leqq G$.

2. *Transformation of a Hermitian Matrix.* Let us suppose then that A is a Hermitian matrix of order n. Denote by A_τ the principal minor matrix of order τ standing in the upper left hand corner of A and by $L_\tau(\lambda) = 0$ the characteristic equation of A_τ. If $\rho_1 \leqq \cdots \leqq \rho_\tau$ are the characteristic roots of A_τ there exists a unitary matrix $P = (p_{ij})$ such that $\bar{P}'A_\tau P = B_\tau$, where B_τ has as elements in its main diagonal the roots $\rho_1, \cdots, \rho_\tau$ and zeros elsewhere. If $A_{\tau+1}$ be the Hermitian matrix of order $\tau + 1$ formed by adjoining to A_τ an additional row $x_1, \cdots, x_\tau, x_{\tau+1}$ ($x_{\tau+1}$ real), and a column consisting of the conjugates of these elements, and if R be the unitary matrix

$$R = \begin{pmatrix} P, & 0 \\ 0, & 1 \end{pmatrix}$$

formed by adjoining to P an additional row and column consisting entirely of zeros except in the last place, it is easy to verify that $\bar{R}'A_{\tau+1}R = B_{\tau+1}$, $B_{\tau+1}$ being a matrix of the form:

(1) $$B_{\tau+1} = \begin{pmatrix} \rho_1, \cdots, 0, \bar{X}_\tau \\ \cdots\cdots\cdots\cdots \\ 0, \cdots, \rho_\tau, \bar{X}_1 \\ X_1, \cdots, X_\tau, X_{\tau+1} \end{pmatrix}$$

where

(2) $$X_{\tau+1} = x_{\tau+1}; \quad X_j = \sum_{i=1}^{\tau} p_{ij}x_i \qquad (j = 1, \cdots, \tau).$$

Under such a transformation the characteristic equations $L_\tau(\lambda) = 0$ and $L_{\tau+1}(\lambda) = 0$ of A_τ and $A_{\tau+1}$ are unaltered.

Expanding the characteristic determinant of $B_{\tau+1}$ according to the elements of its last row and last column, $L_{\tau+1}(\lambda)$ may be written

(3) $$L_{\tau+1}(\lambda) = -\sum_{i=1}^{\tau} X_i\bar{X}_iR_i(\lambda) + (x_{\tau+1} - \lambda)L_\tau(\lambda),$$

where the $R_i(\lambda)$ are defined by the relations

(4) $$(\rho_i - \lambda)R_i(\lambda) = (\rho_1 - \lambda)\cdots(\rho_\tau - \lambda) = L_\tau(\lambda)$$

and are therefore real. Manifestly $R_i(\rho_j) = 0$ $(i \neq j)$ while if the ρ's are all distinct $R_i(\rho_i) \neq 0$.

---

* "The Characteristic Equation of a Matrix," *Bulletin of the American Mathematical Society*, Vol. 34 (1928), pp. 363-368.

3. *The Vanishing of Certain X's.* Since P is nonsingular $(X_1, \cdots, X_\tau) = (0, \cdots, 0)$ if, and only if, $(x_1, \cdots, x_\tau) = (0, \cdots, 0)$. Let us suppose then that the X's are not all zero but that $X_{\gamma+1}, \cdots, X_{\gamma+m}$ which correspond in (1) to a root $\rho_{\gamma+1} = \cdots = \rho_{\gamma+m} = \rho$ (say) of multiplicity m of A_τ are all zero. We then have

$$\sum_{i=1}^{\tau} p_{ij} x_i = 0 \qquad (j = \gamma + 1, \cdots, \gamma + m) \tag{5}$$

so that the set $(x_1, \cdots, x_\tau)$ is a solution of the system of homogeneous linear equations (5) whose coefficients are the $(\gamma + 1)$th, $\cdots$, $(\gamma + m)$th columns of P. But from the manner in which P was built up * we have also the following

$$\sum_{j=1}^{\tau} (a_{ij} - \rho\delta_{ij}) p_{jr} = 0 \qquad (i = 1, \cdots, \tau), \ (r = \gamma + 1, \cdots, \gamma + m) \tag{6}$$

where δ_{ij} is the Kronecker symbol and is equal to 1 if $i = j$; 0 if $i \neq j$. Thus the $\tau - m$ linearly independent rows of $A_\tau - \rho I$ are also solutions of (5), and since the latter system has at most $\tau - m$ linearly independent solutions, it follows that the set $(x_1, \cdots, x_\tau)$ depends linearly on the rows of $A_\tau - \rho I$. Conversely, if $(x_1, \cdots, x_\tau)$ depends linearly on the rows of $A_\tau - \rho I$, $X_{\gamma+1} = \cdots = X_{\gamma+m} = 0$. We therefore have the following theorem:

THEOREM I. *If ρ is a characteristic root of multiplicity m of A_τ and if $X_{\gamma+1}, \cdots, X_{\gamma+m}$ are X's corresponding to this multiple root in the matrix $B_{\tau+1}$; then $X_{\gamma+1} = \cdots = X_{\gamma+m} = 0$, if, and only if, the bordering set $x_1, \cdots, x_\tau$ depends linearly on the rows of $A_\tau - \rho I$.*

4. *The Separation Property.* Let us now suppose that in (3) all the ρ's are distinct and none of the X's is zero. Since $L_{\tau+1}(\lambda) = (-\lambda)^{\tau+1} + \cdots$, manifestly $L_{\tau+1}(-\infty) > 0$ whether τ is even or odd. Also

$$\begin{aligned} L_{\tau+1}(\rho_i) &= -X_i \bar{X}_i R_i(\rho_i) \\ &= -X_i \bar{X}_i (\rho_1 - \rho_i) \cdots (\rho_{i-1} - \rho_i)(\rho_{i+1} - \rho_i) \cdots (\rho_\tau - \rho_i) \\ &= (-1)^i k_i \quad \text{where } k_i > 0 \qquad\qquad (i = 1, \cdots, \tau). \end{aligned} \tag{7}$$

Further, $L_{\tau+1}(\infty)$ has the same sign as $(-1)^{\tau+1}$. Using for uniformity the notation $\rho_0 = -\infty$, $\rho_{\tau+1} = \infty$, we may say that (7) holds also for $i = 0$ and $i = \tau + 1$. It is clear then that in each of the open intervals (ρ_{i-1}, ρ_i) $(i = 1, \cdots, \tau + 1)$ there is exactly one root σ_i of $L_{\tau+1}(\lambda) = 0$. We therefore have

* Kowalewski, *loc. cit.*, pp. 195-196.

THEOREM II. *If the characteristic roots $\rho_1, \cdots, \rho_\tau$ of a Hermitian matrix A_τ are all distinct and if $A_{\tau+1}$ is the Hermitian matrix formed by adjoining to A_τ a row $x_1, \cdots, x_\tau, x_{\tau+1}$ ($x_{\tau+1}$ real)* and *a column $\bar{x}_1, \cdots, \bar{x}_\tau, x_{\tau+1}$, then if the set $x_1, \cdots, x_\tau$ does not depend linearly on the rows of any of the matrices $A_\tau - \rho_i I$ $(i = 1, \cdots, \tau)$ the $\tau + 1$ characteristic roots $\sigma_1, \cdots, \sigma_{\tau+1}$ of $A_{\tau+1}$ are all distinct and are separated by the ρ's.*

Suppose, however, that $\rho_{\gamma+1} = \cdots = \rho_{\gamma+e} = \rho$ is a root of multiplicity e of $L_\tau(\lambda) = 0$. Then evidently each $R_i(\lambda)$ is divisible by $(\rho - \lambda)^{e-1}$ while $R_i(\lambda)$ $(i = \gamma + 1, \cdots, \gamma + e)$ are not divisible by $(\rho - \lambda)^e$. Writing $R_i(\lambda) = (\rho - \lambda)^{e-1} S_i(\lambda)$ and noting that

$$S_{\gamma+1}(\lambda) = \cdots = S_{\gamma+e}(\lambda) = S_\rho(\lambda), \quad \text{say,}$$

we may write $L_{\tau+1}(\lambda)$ in the form

$$L_{\tau+1}(\lambda) = (\rho - \lambda)^{e-1} f(\lambda)$$

where $f(\lambda)$ is an expression of the type (3) with the root ρ now playing the role of a simple root and with the coefficient $X_\rho \bar{X}_\rho$ of $S_\rho(\lambda)$ satisfying the condition

$$X_\rho \bar{X}_\rho = \sum_{i=\gamma+1}^{\gamma+e} X_i \bar{X}_i.$$

Evidently $X_\rho = 0$ if, and only if, the set $x_1, \cdots, x_\tau$ depends linearly on the rows of $A_\tau - \rho I$.

If now $L_\tau(\lambda) = 0$ has the $m \leqq \tau$ distinct roots $\rho_1 < \cdots < \rho_m$ of multiplicities $e_1, \cdots, e_m$, respectively, we may proceed with regard to each of these roots as we did with regard to ρ until finally $L_{\tau+1}(\lambda)$ may be written in the form

$$L_{\tau+1}(\lambda) = (\rho_1 - \lambda)^{e_1 - 1} \cdots (\rho_m - \lambda)^{e_m - 1} F(\lambda)$$

where $F(\lambda)$ is an expression of the type (3) with each ρ playing the role of a simple root. If the set $x_1, \cdots, x_\tau$ does not depend linearly on the rows of any of the matrices $A_\tau - \rho_i I$ all of the X's entering $F(\lambda)$ are different from zero. Hence, writing $\rho_0 = -\infty$, $\rho_{m+1} = \infty$ it follows that $F(\lambda) = 0$ has exactly one root in each of the open intervals (ρ_{i-1}, ρ_i) $(i = 1, \cdots, m + 1)$.

We have therefore proved

THEOREM III. *If $L_\tau(\lambda) = 0$ has the m distinct roots $\rho_1 < \cdots < \rho_m$ of multiplicities $e_1, \cdots, e_m$, respectively, and if the bordering set $x_1, \cdots, x_\tau$ does not depend linearly on the rows of any of the matrices $A_\tau - \rho_i I$ $(i = 1, \cdots, m)$, then each ρ_i is a root of $L_{\tau+1}(\lambda) = 0$ of multiplicity exactly $e_i - 1$, while in each of the open intervals (ρ_{i-1}, ρ_i) $(i = 1, \cdots, m + 1)$ there lies exactly one root of $L_{\tau+1}(\lambda) = 0$.*

Suppose now that $\rho_{\gamma+1} = \cdots = \rho_{\gamma+m} = \rho$ is a root of multiplicity m of $L_\tau(\lambda) = 0$ and that the set $x_1, \cdots, x_\tau$ depends linearly on the rows of $A_\tau - \rho I$, so that $(X_{\gamma+1}, \cdots, X_{\gamma+m}) = (0, \cdots, 0)$. From the determinantal form of $L_{\tau+1}(\lambda)$ it is manifest that the latter contains $(\rho - \lambda)^m$ as a factor, so that ρ is a root of $L_{\tau+1}(\lambda) = 0$ of multiplicity at least m. Indeed, if

$$x_j = \sum_{i=1}^{\tau} (a_{ij} - \rho\delta_{ij}) c_i = \sum_{i=1}^{\tau} a_{ij} c_i - \rho c_j$$

it follows from an examination of the rank of $B_{\tau+1}$ that ρ will be a root of multiplicity $m + 1$ or m of $L_{\tau+1}(\lambda) = 0$ according as $x_{\tau+1}$ is or is not equal to

$$\rho(1 - \sum_{j=1}^{\tau} c_j \bar{c}_j) + \sum_{i,j}^{1,\ldots,\tau} a_{ij} c_i \bar{c}_j.$$

Hence we have the following theorem:

THEOREM IV. *A root ρ of multiplicity m of $L_\tau(\lambda) = 0$ will be a root of multiplicity at least m (and at most $m + 1$) of $L_{\tau+1}(\lambda) = 0$ if, and only if, the bordering set $x_1, \cdots, x_\tau$ depends linearly on the rows of $A_\tau - \rho I$.*

5. *Number of Negative and of Positive Roots of $A_{\tau+1}$.* Let the ν distinct negative roots of $L_\tau(\lambda) = 0$ be $\rho_1 < \cdots < \rho_\nu$ of multiplicities $e_1, e_2, \cdots, e_\nu$, respectively. If the X's corresponding to the roots $\rho_{i_1}, \cdots, \rho_{i_\gamma}$ are all zero the latter are roots of $L_{\tau+1}(\lambda) = 0$ of multiplicities at least $e_{i_1}, \cdots, e_{i_\gamma}$, respectively. If for the remaining ρ's, $\rho_{i_{\gamma+1}}, \cdots, \rho_{i_\nu}$ the corresponding X's are not all zero, these are roots of $L_{\tau+1}(\lambda) = 0$ of multiplicities exactly $e_{i_{\gamma+1}} - 1, \cdots, e_{i_\nu} - 1$, respectively, while in each of the $\nu - \gamma$ open intervals

$$-\infty, \quad \rho_{i_{\gamma+1}} \cdots, \rho_{i_\nu}, \tag{8}$$

there is exactly one (negative) root of $L_{\tau+1}(\lambda) = 0$. Hence, the latter equation has at least as many negative roots as $L_\tau(\lambda) = 0$. Similarly, $L_{\tau+1}(\lambda) = 0$ has at least as many positive roots as $L_\tau(\lambda) = 0$.

If zero is a root of multiplicity m of $L_\tau(\lambda) = 0$ and is likewise a root of multiplicity at least m of $L_{\tau+1}(\lambda) = 0$, it is clear that the latter equation can have at most one more negative (positive) root than the former; while if zero is a root of multiplicity exactly $m - 1$ of $L_{\tau+1}(\lambda) = 0$ by adjoining 0 to the sequence (8) it follows that the latter equation has exactly one more negative root and likewise one more positive root than $L_\tau(\lambda) = 0$.

We may state the theorem as follows:

THEOREM V. *If m, ν and μ represent the numbers of zero, negative and positive roots of $L_\tau(\lambda) = 0$ and if Z, N and P represent the corresponding numbers for $L_{\tau+1}(\lambda) = 0$, then, if $Z = m - 1$, $N = \nu + 1$* and *$P = \mu + 1$;*

if $Z=m$, $N=\nu+1$ or ν, $P=\mu$ *or* $\mu+1$; *and finally, if* $Z=m+1$, $N=\nu$ *and* $P=\mu$.

6. *The Signature of a Hermitian Matrix.* If $L_\tau(\lambda)=0$ has ν negative roots and μ positive roots, the difference $\mu-\nu$ is called the *signature* * of A_τ. Denote by M_i the determinant of the matrix A_i. Suppose now that A_τ is non-singular, i. e., $M_\tau \neq 0$. If $A_{\tau+1}$ is also non-singular, by Theorem V $L_{\tau+1}(\lambda)=0$ will have ν negative and $\mu+1$ positive roots or $\nu+1$ negative and μ positive roots according as M_τ and $M_{\tau+1}$ have the same sign or opposite signs. That is, the signature of $A_{\tau+1}$ is greater or less by one than the signature of A_τ according as the sequence of two terms M_τ, $M_{\tau+1}$ presents a permanence or a variation of sign.

But if $A_{\tau+1}$ is singular and therefore $L_{\tau+1}(\lambda)=0$ has one zero root, the latter has exactly ν negative and μ positive roots. If further $A_{\tau+2}$ is non-singular, $L_{\tau+2}(\lambda)=0$ has by Theorem V exactly $\nu+1$ negative and $\mu+1$ positive roots. Hence, $M_{\tau+2}$ is of opposite sign to M_τ. Moreover, the signatures of $A_{\tau+2}$ and A_τ are the same. Noting that the matrix $\begin{pmatrix} 0, & a_{12} \\ \bar{a}_{12}, & 0 \end{pmatrix}$ $(a_{12}\neq 0)$ has one negative and one positive characteristic root, it is clear that we have established Gundelfinger's † rule for determining the signature of a regularly arranged Hermitian or real symmetric matrix.

THEOREM VI. *If a Hermitian or a real symmetric matrix of rank r is regularly arranged, i. e., if the rows and columns are so arranged that no two consecutive terms in the sequence*

$$(9) \qquad M_0=1, \quad M_1=a_{11}, \cdots, M_r=\begin{vmatrix} a_{11}, \cdots, a_{1r} \\ \cdots\cdots \\ a_{r1}, \cdots, a_{rr} \end{vmatrix},$$

are zero and $M_r\neq 0$, *the signature of the matrix is equal to the difference between the number of permanences of sign and the number of variations of sign in the sequence* (9), *where a vanishing term may be counted as either positive or negative, but must be counted.*

7. *Application to Hermitian Matrices which are not Regularly Arranged.* Suppose now that both $A_{\tau+1}$ and $A_{\tau+2}$ are singular while A_τ is not. Let us denote by Z, N and P the numbers of zero, negative and positive roots of an equation under consideration. It is clear that if for $L_\tau(\lambda)=0$ we have

$$L_\tau(\lambda): \qquad Z=0, \quad N=\nu, \quad P=\mu,$$

* μ is sometimes called the index of A; cf. Dickson, *loc. cit.*, p. 71.

† Gundelfinger, "Zur Theorie der quadratischen Formen," *Crelle*, Vol. 91 (1881), p. 225; cf. also Dickson, *loc. cit.*, pp. 87-88.

then for $L_{\tau+1}(\lambda)=0$ we have

$$L_{\tau+1}(\lambda): \quad Z=1, \quad N=\nu, \quad P=\mu,$$

and for $L_{\tau+2}(\lambda)=0$ we have one of the following

$$(10) \qquad L_{\tau+2}(\lambda): \quad \begin{array}{lll} Z=2, & N=\nu, & P=\mu; \\ Z=1, & N=\nu+1, & P=\mu; \\ Z=1, & N=\nu, & P=\mu+1. \end{array}$$

If $A_{\tau+3}$ is non-singular (so that for $L_{\tau+2}(\lambda)$ the case $Z=2$ cannot arise), the outlay for $L_{\tau+3}(\lambda)$ is by Theorem V

$$(11) \qquad L_{\tau+3}(\lambda): \quad \begin{array}{lll} Z=0, & N=\nu+2, & P=\mu+1; \\ Z=0, & N=\nu+1, & P=\mu+2. \end{array}$$

That is, if $M_\tau M_{\tau+3} \neq 0$ while $M_{\tau+1}=M_{\tau+2}=0$, $L_{\tau+3}(\lambda)=0$ has two more or one more negative roots than $L_\tau(\lambda)=0$ according as M_τ and $M_{\tau+3}$ have the same sign or opposite signs.

Suppose, however, that $A_{\tau+1}$, $A_{\tau+2}$ and $A_{\tau+3}$ are singular while A_τ and $A_{\tau+4}$ are not. The possibilities for $L_{\tau+3}(\lambda)=0$ are then easily seen to be:

$$(12) \qquad L_{\tau+3}(\lambda): \quad \begin{array}{lll} Z=1, & N=\nu+1, & P=\mu+1; \\ Z=1, & N=\nu+2, & P=\mu; \\ Z=1, & N=\nu, & P=\mu+2; \end{array}$$

and for $L_{\tau+4}(\lambda)=0$:

$$(13) \qquad L_{\tau+4}(\lambda): \quad \begin{array}{lll} Z=0, & N=\nu+2, & P=\mu+2; \\ Z=0, & N=\nu+3, & P=\mu+1; \\ Z=0, & N=\nu+1, & P=\mu+3. \end{array}$$

If M_τ and $M_{\tau+4}$ are of the same sign, manifestly the first case in (13) is the only one that can arise, while if M_τ and $M_{\tau+4}$ are of opposite signs, either of the last two cases may arise and we cannot distinguish between them by the signs of the M's alone.

We therefore have the Theorem:

THEOREM VII. *If in the sequence* (9) $M_\tau \neq 0$ *and* $M_\tau M_{\tau+3} \neq 0$ *while* $M_{\tau+1}=M_{\tau+2}=0$, *then to the subsequence* M_τ, 0, 0, $M_{\tau+3}$ *we assign two variations and one permanence or one variation and two permanences of sign according as* M_τ *and* $M_{\tau+3}$ *have the same sign or opposite signs; and if* $M_\tau M_{\tau+4} \neq 0$ *while* $M_{\tau+1}=M_{\tau+2}=M_{\tau+3}=0$ *we assign to the subsequence* M_τ, 0, 0, 0, $M_{\tau+4}$ *exactly two variations and two permanences if* M_τ *and* $M_{\tau+4}$ *have the same sign, while in the contrary case the number of variations to be assigned may be either one or three.*

While the last theorem was proved only on the supposition that $M_\tau \neq 0$ for $\tau > 0$ it is easy to verify that the results hold also for $\tau = 0$.

The questions discussed in this section were studied originally by Frobenius,* and when two consecutive terms in the sequence (9) vanish the results that he arrived at by a very elaborate discussion are exactly the results that we have arrived at here. When three consecutive terms vanish and the adjacent M's have *opposite* signs, Frobenius points out that the signature of the matrix is not determined by the sequence (9) alone. But he does not seem to show that the signature *is* definitely determined when the adjacent M's have the *same* sign. More recently Franklin † attacked the same problem by a scheme similar to, but, it seems to the author, less explicit and less powerful than, the one used here, and he arrived at the same conclusions that Frobenius had previously arrived at. Still more recently and by an entirely different method the author ‡ obtained the results here given.

8. *A Sequence of Sturm Functions for the Equation* $L_n(\lambda)=0$. Let α and β be any two real numbers, neither a root of $L_n(\lambda)=0$. Since the characteristic roots of $A-\alpha I$ are less by α than the characteristic roots of A, it is clear that if ν_α is the number of characteristic roots $<\alpha$ of A, then ν_α is the number of negative roots of $A-\alpha I$. If in the sequence

$$(14) \qquad 1,\quad L_1(\alpha),\quad L_2(\alpha),\cdots,L_n(\alpha)$$

not more than two consecutive terms vanish (or if three consecutive terms vanish and the adjacent terms have the same sign), ν_α is equal to the number of variations of sign in the sequence, where if two or more consecutive terms vanish the number of variations is determined by theorem VII. Here a root of multiplicity m counts as m roots. Under the same restrictions if ν_β is the number of variations of sign in the sequence (14) with α replaced by β, then ν_β is the number of roots $<\beta$ of $L_n(\lambda)=0$. Hence for $\alpha<\beta$ $\nu_\beta-\nu_\alpha$ is the number of roots of $L_n(\lambda)=0$ between α and β. Without altering the number of variations of sign the order of the terms in (14) may be reversed thus furnishing a sequence in which the last one is always greater than zero. Such a sequence therefore

$$L_n(\lambda),\ L_{n-1}(\lambda),\cdots,L_1(\lambda),\ 1$$

may be thought of as constituting a sequence of Sturm functions § for the equation $L_n(\lambda)=0$.

THE UNIVERSITY OF NORTH CAROLINA.

---

* Frobenius, "Ueber das Trägsheitsgesetz der quadratischen Formen," *Crelle*, Vol. 114 (1895), pp. 198-199.

† Franklin, "A Theorem of Frobenius on Quadratic Forms," *Bulletin of the American Mathematical Society*, Vol. 33 (1927), pp. 451-452.

‡ "On the Signature of a Quadratic Form," *Annals of Mathematics*, 2nd Series, Vol. 30 (1929), pp. 517-525.

§ Cf. Salmon, *Lessons on Higher Algebra*, Third Edition, Dublin (1876), p. 43.

Discontinuous Solutions in the Problem of Depreciation and Replacement.

By Henry H. Pixley.

1. *Introduction.* The mathematics of the problem of depreciation in economics has been the subject of recent papers by Hotelling * and by Roos.† Roos has developed a dynamical theory of depreciation and replacement and has formulated the problem of replacement for a single operating machine as a type of Lagrange problem in the calculus of variations. The expression which he maximizes is the sum of two definite integrals whose integrands are functions of variable end and corner values. He considers it as a single integral with an integrand which is discontinuous along a continuous curve of corners. The maximizing arc which he obtains is, however, continuous at the time of replacement. This means that the replacement machine starts at the time and at the rate of production at which the operating machine stops. In an actual case this is not necessarily true.

In this paper I develop a general theory corresponding to that of Roos without the assumptions of continuity at the time of replacement. In particular, an application is given in which the replacement machine is started at a time and rate different from those at which the operating machine stops.

2. *The replacement problem.* We consider a situation in which one machine operates from time t_1 to time w_1 at the rate of $u_1(t)$ units of output per unit time. Of the output of the machine $y_1(t)$ units are sold per unit time at a price $p_1(t)$ per unit. The total operating cost of the machine including depreciation is represented by the function $Q_1(u_1, u_1', p_1, p_1', t)$. A second machine operates from time $w_2(\geqq w_1)$ to time t_2 with an output of $u_2(t)$ and a demand of $y_2(t)$ which sells at $p_2(t)$ per unit. The corresponding cost function is $Q_2(u_2, u_2', p_2, p_2', t)$.

Roos has shown that the total value, discounted to the time T, of the

* H. Hotelling, "A General Mathematical Theory of Depreciation," *Journal of American Statistical Association* (September, 1925).

† C. F. Roos, "A Mathematical Theory of Depreciation and Replacement," *American Journal of Mathematics*, Vol. 50 (January, 1928); Roos, "The Problem of Depreciation in the Calculus of Variations," *Bulletin of the American Mathematical Society*, Vol. 34 (1928), p. 218.

profits from a machine which operates from T_1 to T_2 plus the value at T of its scrap value at T_2 is

$$V(T) = \int_{T_1}^{T_2} [p(t)y(t) - Q(u, u', p, p', t)]\, E(T, t)\, dt + KE(T, T_2)$$

where K is the initial cost of the machine, and $E(T, t)$ is $\exp\left[-\int_T^t \delta(\nu)\, d\nu\right]$ in which $\delta(\nu)$ is the rate of increase of an invested sum s divided by s. The function $E(T, t)$ is a discount (or interest) factor which gives the value at T of the profits earned at t.* By this formula the total value at the time T of the two machines minus the value at T of the amounts necessary to replace the machines at w_2 and t_2 respectively is

$$\begin{aligned} & V_1(T) - K_1 E(T, w_2) + V_2(T) - K_2 E(T, t_2) \\ (1) \qquad & = \int_{t_1}^{w_1} (p_1 y_1 - Q_1) E(T, t)\, dt + K_1 [E(T, w_1) - E(T, w_2)] \\ & \quad + \int_{w_1}^{w_2} (p_0 y_0 - Q_0) E(T, t)\, dt + \int_{w_2}^{t_2} (p_2 y_2 - Q_2) E(T, t)\, dt \end{aligned}$$

where the subscripts 1 and 2 denote functions of the operating and replacement machines respectively and the subscript 0 denotes functions for the period of replacement from w_1 to w_2. The cost function Q_0 represents any variable expense which occurs during the period of replacement and which may not be considered part of the constant K_2. The function Q_0 may be a function of w_1 and w_2.

For convenience we will drop the subscripts 1 and 2 for the present and let $u(t)$, $y(t)$, $p(t)$, and $Q(t)$ represent the rate of output, rate of demand, price, and cost of production, respectively, for the range $t_1 \leqq t \leqq t_2$. These functions may be discontinuous for the values $t = w_1$, w_2 but are continuous for all other values of t in the range $t_1 \leqq t \leqq t_2$. The functions $u(t)$, $y(t)$, and $p(t)$ are not in general independent, since y and p are related by an equation of demand, while y and u satisfy an equation of supply. If we assume that the demand equation is a first order differential equation of the form $\theta(y, y', p, p', t) = 0$, and that the supply equation is $y = \xi(u, t)$, we can obtain by the elimination of y, a demand-supply equation †

$$(2) \qquad \phi(u, u', p, p', t) = 0.$$

---

* Roos, "The Problem of Depreciation in the Calculus of Variations," *loc. cit.*, p. 221.

† Roos, "A Dynamical Theory of Economics," *The Journal of Political Economy*, Vol. 35 (October, 1927); Roos, "The Problem of Depreciation in the Calculus of Variations," *loc. cit.*, p. 222.

There will also in general be certain conditions which the end-points must satisfy which may be written in the form

$$(3) \qquad \psi_\mu\,[t_\sigma, u(t_\sigma), p(t_\sigma), w_\sigma, u(w_\sigma), p(w_\sigma \pm 0)] = 0, \qquad (\mu = 1, \cdots, p \leqq 14),$$

where σ may take both of the values 1, 2 in each of the p equations.

We will eliminate y from the function (1) by means of the equation $y = \xi(u, t)$. Then if we assume that this function is to be maximized, our problem is that of finding among the arcs $u(t)$, $p(t)$, satisfying the equation (2), and whose end-points satisfy equations (3), a set which maximizes this expression (1).

3. *The general problem.* We will now state a more general problem of which the problem of the preceding paragraph is a special case. We will need to consider a class of arcs, $y_i = y_i(x)$, $(i = 1, \cdots, n)$, which are defined for $x = s_1$, $x = s_2$, and $x = s_3$, where $x_1 \leqq s_1 \leqq x_2$, $x_3 \leqq s_2 \leqq x_4$, $x_5 \leqq s_3 \leqq x_6$. We will represent these three intervals by the letters X_1, X_2, and X_3, respectively. Our general problem is that of *finding among those discontinuous arcs, $y_i = y_i(x)$, $(i = 1, \cdots, n)$, of the above class which satisfy certain differential equations*

$$(4) \qquad \phi_\alpha(x, y, y') = 0, \qquad (\alpha = 1, \cdots, m < n),$$

for all x in X_1, X_2, and X_3, and whose points at the ends of these intervals satisfy the end equations

$$(5) \qquad \psi_\mu\,[x_\rho, y(x_\rho)] = 0, \qquad (\mu = 1, \cdots, p \leqq 6n + 6;\ \rho = 1, \cdots, 6),$$

one which maximizes an expression

$$(6) \qquad I = \int_{x_1}^{x_2} f\,[s_1, y, y', x_\rho, y(x_\rho)]\,ds_1 + \int_{x_3}^{x_4} g\,[s_2, y, y', x_\rho, (x_\rho)]\,ds_2 + \int_{x_5}^{x_6} h\,[s_3, y, y', x_\rho, y(x_\rho)]\,ds_3.$$

where (y, y') represents the set $(y_1, \cdots, y_n, y_1', \cdots, y_n')$, $[x_\rho, y(x_\rho)]$ represents the set

$$[x_1, y_1(x_1), \cdots, y_n(x_1), x_2, y_1(x_2), \cdots, y_n(x_2), \cdots, x_6, y_1(x_6), \cdots, y_n(x_6)],$$

and primes denote differentiation with respect to x.

We assume that:

(a) the functions $y_i(x)$ defining the maximizing arc E are continuous

in each of the intervals X_1, X_2, X_3, and have continuous derivatives in these intervals except at a finite number of values of x;

(b) in a neighborhood R of the values (x, y, y') on the arc E the functions f, g, h, and ϕ_α have continuous derivatives up to and including those of the second order;

(c) at every element (x, y, y') on E the $m \times n$-dimensional matrix $\|\phi_{\alpha y_i'}\|$ has rank m;

(d) the functions ψ_μ have continuous derivatives up to and including those of the second order near the end values $[x_\rho, y(x_\rho)]$ and at these values the $p \times (6n+6)$-dimensional matrix

$$\| \psi_{\mu x_\rho} \quad \psi_{\mu y_\rho} \| \tag{7}$$

has rank p, where $y_\rho = y(x_\rho)$, $(\rho = 1, \cdots, 6)$, and the subscripts y_i', x_ρ, y_ρ denote partial derivatives.*, †

For the general problem as here stated certain necessary conditions for a solution can be obtained by methods which are essentially those given by Bliss for the problem with a continuous integrand,* and which have been extended by Roos to the case of a discontinuous integrand.† In each of these treatments the solution is sought in a class of continuous arcs.

4. *Admissible arcs and variations.* An arc $y_i = y_i(x)$, $(i = 1, \cdots, n)$, defined over the intervals X_1, X_2, X_3 will be called an *admissible arc* if it has the continuity properties (a); if all of its elements (x, y, y') lie in R, and if it satisfies the differential equations (4).

If a one-parameter family of admissible arcs

$$y_i = y_i(x, b), \tag{8}$$

$[i = 1, \cdots, n;\ x_1(b) \leqq x \leqq x_2(b);\ x_3(b) \leqq x \leqq x_4(b);\ x_5(b) \leqq x \leqq x_6(b)]$ containing a particular admissible arc E for the parameter value $b = 0$ be given, the functions $\eta_i(x) = \partial y_i(x, 0)/\partial b$, $\xi_\rho = \partial x_\rho(0)/\partial b$ are called *variations of the family along E*.

The equations of variation on the arc E for the functions ϕ_α are defined by

$$\Phi_\alpha(\eta, \eta') = \phi_{\alpha y_i}\eta_i + \phi_{\alpha y_i'}\eta_i' = 0, \qquad (\alpha = 1, \cdots, m), \tag{9}$$

* G. A. Bliss, "Lectures on the Problem of Lagrange in the Calculus of Variations," *University of Chicago* (1925), mimeographed by O. E. Brown, University of Chicago.

† Roos, "General Problem of Minimizing an Integral with Discontinuous Integrand," *Transactions of the American Mathematical Society*, Vol. 31, (January, 1929), (hereafter referred to as "General Problem").

where the coefficients ϕ_{ay_i}, $\phi_{ay_i'}$ have as arguments the functions $y_i(x)$ defining the arc E and the functions η_i, η_i' are, of course, defined only for values of x in the intervals X_1, X_2, X_3.

Similarly we define the equations of variation on the arc E for the functions ψ_μ to be

$$\Psi_\mu(\xi,\eta)=\psi_{\mu x_\rho}\xi_\rho+\psi_{\mu y_{i\rho}}dy_i[x_\rho(0),0]/db, \tag{10}$$

where in equations (9) and (10) i is an umbral index with range $1,\cdots,n$, and ρ is umbral with range $1,\cdots,6$, according to the convention that whenever a subscript appears twice in a term that term is to be summed for all values of the subscript. The functions Ψ_μ are clearly functions of ξ_ρ and η_i since

$$dy_i[x_\rho(0),0]/db=y_{i\rho}'\xi_\rho+\eta_i(x_\rho),\quad(\rho=1,\cdots,6,\text{ not umbral}). \tag{11}$$

A set of arbitrary constants ξ_ρ and functions $\eta_i(x)$ with the continuity properties described in (a) and satisfying the equations of variation (9) will be called a *set of admissible variations,* a definition which we will find useful since

For every set of admissible variations ξ_ρ, $\eta_i(x)$ along the arc E there exists a one-parameter family (8) *of admissible arcs containing E for the value $b=0$ and having the set ξ_ρ, $\eta_i(x)$ as its variations along E. For this family the functions $y_i(x,b)$ are continuous on each of the intervals X_1, X_2, X_3 and have continuous derivatives with respect to b for all values (x,b) near those defining E, and the derivatives $dy_i(x,b)/dx$ have the same property except, possibly, at the values of x defining corners of E.**

5. *First necessary conditions.* If we substitute the one-parameter family of admissible arcs, (8), containing E for $b=0$, in the expression I, differentiate I with respect to b, and set $b=0$, we obtain the first variation of I along the arc E

$$\begin{aligned}I_1(\xi,\eta)=&\int_{x_1}^{x_2}(f_{y_i}\eta_i+f_{y_i'}\eta_i')\,ds_1\\&+\int_{x_3}^{x_4}(g_{y_i}\eta_i+g_{y_i'}\eta_i')\,ds_2+\int_{x_5}^{x_6}(h_{y_i}\eta_i+h_{y_i'}\eta_i')\,ds_3\\&+K_{i\rho}(f,g,h)\,dy_i[x_\rho(0),0]/db+L_\rho(f,g,h)\xi_\rho,\end{aligned} \tag{12}$$

---

* For proof see Roos, "General Problem," *loc. cit.*, p. 61. See also Bliss, "Lectures, etc.," *loc. cit.*, p. 4. The theorem stated above is an obvious extension of the one proved by Roos.

where $$K_{i\rho}(f,g,h) = \int_{x_1}^{x_2} f_{y_i\rho} ds_1 + \int_{x_3}^{x_4} g_{y_i\rho} ds_2 + \int_{x_5}^{x_6} h_{y_i\rho} ds_3;$$
$$L_\rho(f,g,h) = f_\rho + \int_{x_1}^{x_2} f_{x\rho} ds_1 + \int_{x_3}^{x_4} g_{x\rho} ds_2 + \int_{x_5}^{x_6} h_{x\rho} ds_3; \text{*}$$

$f_1 = -f(x_1)$, $f_2 = f(x_2)$, $f_3 = -g(x_3)$, $f_4 = g(x_4)$, $f_5 = -h(x_5)$, $f_6 = h(x_6)$; $f(x_1)$ is the value of the function at the end-point of the arc E corresponding to $x = x_1$ and the other functions, f_ρ, are similarly defined; i is an umbral index with range $1, \cdots, n$, and ρ is umbral with range $1, \cdots, 6$; and the subscripts y_i, y_i', $y_{i\rho}$, x_ρ denote partial derivatives.

Following the methods of Bliss and Roos it can be proved by means of this first variation that: *For every maximizing arc for the above problem there exist sets of constants* c_{i1}, c_{i2}, c_{i3}, $(i = 1, \cdots, n)$, *and functions*

$$F(s_1, y, y', x_\rho, y_{i\rho}, \lambda_0, \lambda_\alpha) = \lambda_0 f + \lambda_\alpha \phi_\alpha,$$
$$G(s_2, y, y', x_\rho, y_{i\rho}, \lambda_0\ \lambda_\alpha) = \lambda_0 g + \lambda_\alpha \phi_\alpha,$$
$$H(s_3, y, y', x_\rho, y_{i\rho}, \lambda_0\ \lambda_\alpha) = \lambda_0 h + \lambda_\alpha \phi_\alpha, \qquad (\alpha = 1, \cdots, m;\ \text{umbral}),$$

such that the equations

$$F_{y_i'} = \int_{x_1}^{s_1} F_{y_i} ds_1 + c_{i1},\ G_{y_i'} = \int_{x_3}^{s_2} G_{y_i} ds_2 + c_{i2},\ H_{y_i'} = \int_{x_5}^{s_3} H_{y_i} ds_3 + c \tag{13}$$

are satisfied at every point of E. The constant λ_0 and the functions $\lambda_\alpha(x)$, $(\alpha = 1, \cdots, m)$, are not all identically zero on the intervals X_1, X_2, X_3, and are continuous except possibly at values of x defining corners of E. Furthermore, the end values of E must be such that all determinants of order $p+1$ of the matrix

$$\left\| \begin{matrix} N_\rho(x_\nu) & M_{i\rho}(x_\nu) \\ \psi_{\mu x\rho} & \psi_{\mu y_i\rho} \end{matrix} \right\| \tag{14}$$

vanish, where

$$M_{i\rho}(x_\nu) = F_{i\rho} + K_{i\rho}(F,G,H), \quad N_\rho(x_\nu) = -F_{i\rho} y'_{i\rho} + L_\rho(F,G,H);$$

$F_{i1} = -F_{y_i'}(x_1)$, $F_{i2} = F_{y_i'}(x_2)$, $F_{i3} = -G_{y_i'}(x_3)$, $F_{i4} = G_{y_i'}(x_4)$, $F_{i5} = -H_{y_i'}(x_5)$, $F_{i6} = H_{y_i'}(x_6)$; $F_{y_i'}(x_1)$ denotes a derivative with respect to y_i' evaluated at the end-point of E defined by $x = x_1$, and the other functions $F_{i\rho}$ are similarly defined; i is umbral with the range $1, \cdots, n$; $\rho = 1, \cdots, 6$ and ρ is not umbral; and (x_ν) denotes the set $(x_1, \cdots, x_6)$.

* The notation here is suggested by Roos. See "General Problem," *loc. cit.*, p. 62.

6. *The maximizing arcs for the replacement problem.* In the problem stated in § 2 we will assume that the relation, $y = \xi(u, t)$, between the rate of demand and the rate of supply is of the form $y(t) = \alpha_\sigma u(t) + \beta_\sigma(t)$, and furthermore that the demand is a linear function of the price and the rate of change of price, $y(t) = d_\sigma p(t) + e_\sigma(t) + k_\sigma p'(t)$. Then the demand-supply equation, (2), becomes

$$\phi_{1\sigma} = u - a_\sigma p(t) - b_\sigma(t) - h_\sigma p'(t) = 0, \qquad (\sigma = 1, 2), \tag{15}$$

where, as in § 2, $\sigma = 1$, and $\sigma = 2$ denote functions of the operating and replacement machines, respectively; $a_\sigma = d_\sigma/\alpha_\sigma$, $b_\sigma = (e_\sigma + \beta_\sigma)/\alpha_\sigma$, $h_\sigma = k_\sigma/\alpha_\sigma$; and it must be remembered that the forms of the expressions represented by $u(t)$, $y(t)$, and $p(t)$ are not in general the same for the operating and replacement machines. For the period of replacement from the time w_1 to time w_2 we have $u \equiv 0$, and in the place of equation (15) we use the demand equation $y(t) = d_0 p(t) + e_0(t) + k_0 p'(t)$. If in addition we know the initial time, $t_1 = T_1$, the rate and price of output at time t_1, the rate of output at time w_2, and the time which elapses between w_1 and w_2, the conditions (3) are

$$\begin{gathered} \psi_1 = t_1 - T_1 = 0, \quad \psi_2 = u(T_1) - U_1 = 0, \quad \psi_3 = p(T_1) - P_1 = 0, \\ \psi_4 = w_1 - w_2 + W = 0, \quad \psi_5 = u(w_2) - U_2 = 0, \end{gathered} \tag{16}$$

in which T_1, U_1, P_1, W, U_2 are known constants.

Let us also suppose that the cost function Q is expressible by means of the forms

$$\begin{aligned} &Q_\sigma(u, u', p, p', t) \\ &\quad = A_\sigma u^2 + B_\sigma u + C_\sigma + D_\sigma u'^2 + E_\sigma p'^2 + F_\sigma u' + G_\sigma p' + H_\sigma p^2 + I_\sigma p, \\ &\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (\sigma = 1, 2), \\ &Q_0(p, p', t) \\ &\quad = \qquad C_0(t) \qquad + E_0 p'^2 \qquad + G_0 p' + H_0 p^2 + I_0 p. \end{aligned}$$

The parameters a_σ, b_σ, α_σ, β_σ, d_0, e_0, k_0, h_σ, $A_\sigma, \cdots, I_\sigma$, C_0, E_0, G_0, H_0, I_0 are either known functions of the time or constants. In the following solution we will consider all of them except b_σ, C_0, and e_0 as constants for simplicity of the solution, although the problem can be solved when they are functions of t. We will also consider $\delta(\nu)$ a constant.

Our problem may now be stated as that of: *Finding among the arcs* $u(t)$, $p(t)$ *satisfying a demand-supply equation* (15), *and whose end-points satisfy equations* (16), *a set which maximizes the expression* (1), *which may be written*

$$
\begin{aligned}
I = &\int_{t_1}^{w_1} (\alpha_1 p_1 u_1 + \beta_1 p_1 - Q_1) E(T, t)\, dt \\
(17) \quad &+ \int_{w_1}^{w_2} \left[\frac{K_1[E(T, w_1) - E(T, w_2)]}{(w_2 - w_1)} + \{ (d_0 p_0 + e_0 + k_0 p_0') p_0 - Q_0 \} E(T, t) \right] dt \\
&+ \int_{w_2}^{t_2} (\alpha_2 p_2 u_2 + \beta_2 p_2 - Q_2) E(T, t)\, dt.
\end{aligned}
$$

This is a special case of the general problem stated in § 3 where $x_1 = t_1$, $w_1 = x_2 = x_3$, $x_4 = x_5 = w_2$, $x_6 = t_2$, $y_1(x) = u(t)$, $y_2(x) = p(t)$, ϕ_1 is ϕ_{11} for $t_1 \leqq t \leqq w_1$, u for $w_1 \leqq t \leqq w_2$, and ϕ_{12} for $w_2 \leqq t \leqq t_2$, and f, g, and h correspond to the three integrand functions. Therefore the arcs $u(t)$, $p(t)$ with their end-points must satisfy the equations (13) and the tranversality conditions (14).

If we define F, G, and H by the equations *

$$
\begin{aligned}
F = [&\alpha_1 p u + \beta_1 p - A_1 u^2 - B_1 u - C_1 - D_1 u'^2 - E_1 p'^2 - F_1 u' \\
&- G_1 p' - H_1 p^2 - I_1 p + \lambda_{11}(u - a_1 p - h_1 p' - b_1(t))]\, E(T, t) \\
G = K_1&[E(T, w_1) - E(T, w_2)]/(w_2 - w_1) \\
&+ [(d_0 p + e_0 + k p') p - C_0(t) - E_0 p'^2 - G_0 p' - H_0 p^2 - I_0 p] E(T, t) \\
H = [&\alpha_2 p u + \beta_2 p - A_2 u^2 - B_2 u - C_2 - D_2 u'^2 - E_2 p'^2 - F_2 u' \\
&- G_2 p' - H_2 p^2 - I_2 p_2 + \lambda_{12}(u - a_2 p - h_2 p' - b_2(t))]\, E(T, t)
\end{aligned}
$$

we obtain

$$
\begin{aligned}
\partial F/\partial u &= (\alpha_1 p - 2A_1 u - B_1 + \lambda_{11}) E(T, t), \\
\partial F/\partial u' &= (-2D_1 u' - F_1) E(T, t), \\
\partial F/\partial p &= (\alpha_1 u + \beta_1 - 2H_1 p - I_1 - a_1 \lambda_{11}) E(T, t), \\
\partial F/\partial p' &= (-2E_1 p' - G_1 - h_1 \lambda_{11}) E(T, t).
\end{aligned}
$$

The Euler-Lagrange equations in their classical form $dF_{y_i'}/dx = F_{y_i}$ are obtained from equations (13) by differentiation, and in our case these conditions are

$$
(18) \quad
\begin{aligned}
&-2D_1 u'' + (2D_1 u' + F_1)\delta = \alpha_1 p - 2A_1 u - B_1 + \lambda_{11} \\
&-2E_1 p'' - h_1 \lambda_{11}' + (2E_1 p' + G_1 + h_1 \lambda_{11})\delta = \alpha_1 u + \beta_1 - 2H_1 p - I_1 - a_1 \lambda_{11}
\end{aligned}
$$

from each of which the common factor $E(T, t)$ has been removed. Solving the first of these equations for λ_{11} and substituting its value in the second, we obtain

$$
\begin{aligned}
2h_1 D_1 u''' &- 2(a_1 + 2h_1\delta) D_1 u'' + 2(-h_1 A_1 + \gamma_1 \delta D_1) u' + (2\gamma_1 A_1 - \alpha_1) u - 2E_1 \\
&+ (\alpha_1 h_1 + 2\delta E_1) p' + (-\alpha_1 \gamma_1 + 2H_1) p + \gamma_1 (B_1 + \delta F_1) + \delta G_1 + I_1 - \beta_1 =
\end{aligned}
$$

where $\gamma_1 = a_1 + h_1 \delta$.

---

* See Roos, "A Mathematical Theory of Depreciation and Replacement," *loc. cit.*, p. 153.

Replacing u and its derivatives in this equation by their values in terms of p obtained from the demand-supply equation (15), we have the differential equation

$$(19)\quad L_{14}D_t^4p + L_{13}D_t^3p + L_{12}D_t^2p + L_{11}D_tp + L_{10}p + p_{10} = L_1(b_1, b_1', b_1'', b_1''')$$

in which

$$L_{14} = 2h_1^2D_1,\ L_{13} = -4h_1^2\delta D_1,\ L_{12} = -2h_1^2A_1 + 2(-a_1\gamma_1 + h_1^2\delta^2)D_1 - 2E_1,$$

$$L_{11} = 2\delta(h_1^2A_1 + a_1\gamma_1D_1 + E_1),\quad L_{10} = 2a_1\gamma_1A_1 + 2H_1 - 2a_1\alpha_1 - h_1\alpha_1\delta,$$

$$L_1(b_1, b_1', b_1'', b_1''') = -2h_1D_1b_1''' + 2(a_1 + 2h_1\delta)D_1b_1'' + 2(h_1A_1 - \gamma_1\delta D_1)b_1' + (-2\gamma_1A_1 + \alpha_1)b_1,$$

$$p_{10} = \gamma_1(B_1 + \delta F_1) + \delta G_1 + I_1 - \beta_1.$$

Since this is a linear differential equation with constant coefficients its solution depends upon the roots, m_{11}, m_{12}, m_{13}, m_{14}, of the algebraic equation $L_{14}m^4 + L_{13}m^3 + L_{12}m^2 + L_{11}m + L_{10} = 0$. If these roots are all distinct and if $\bar{p}_1(t)$ is a particular solution of equation (19), then its general solution is

$$(20)\qquad p_1 = \bar{p}_1(t) + K_{11}e^{m_{11}t} + K_{12}e^{m_{12}t} + K_{13}e^{m_{13}t} + K_{14}e^{m_{14}t},$$

where the constants K_{11}, K_{12}, K_{13}, K_{14} are arbitrary.

The determination of $\bar{p}_1$ depends, of course, on $b_1(t)$. An interesting form of $b_1(t)$ is the general solution of the homogeneous linear differential equation $L_1(b_1, b_1', b_1'', b_1''') = 0$. The auxiliary equation in this case is $-2h_1D_1\mu^3 + 2(a_1 + 2h_1\delta)D_1\mu^2 + 2(h_1A_1 - \gamma_1\delta D_1)\mu - 2\gamma_1A_1 + \alpha_1 = 0$ and if its roots μ_{11}, μ_{12}, μ_{13} are all distinct the general solution of the equation $L_1 = 0$ is

$$(21)\qquad b_1(t) = \bar{K}_{11}e^{\mu_{11}t} + \bar{K}_{12}e^{\mu_{12}t} + \bar{K}_{13}e^{\mu_{13}t},$$

where $\bar{K}_{11}$, $\bar{K}_{12}$, $\bar{K}_{13}$ are arbitrary constants. The constants $\bar{K}_{11}$, $\bar{K}_{12}$, $\bar{K}_{13}$ are at the disposal of the operator in forming a satisfactory demand-supply equation (15). Hence for this form of $b_1(t)$ our demand-supply equation has five arbitrary constants and at the same time gives us a solution for $p_1(t)$ which can always be expressed explicitly in the form (20). Since $L_1 = 0$, the particular solution may be taken $\bar{p}_1 = -p_{10}/L_{10}$.

It may appear that the price p_1 as given in (20) is independent of the constants $\bar{K}_{11}$, $\bar{K}_{12}$, $\bar{K}_{13}$ in $b_1(t)$. However, in practice, for any change in $\bar{K}_{11}$, $\bar{K}_{12}$, $\bar{K}_{13}$ one would probably choose different values for a_1 and h_1 in the demand-supply equation (15), and p_1 is a function of these constants.

The differential equation which gives $p_2(t)$ is formally like (20), its coefficients being functions of a_2, $b_2(t)$, h_2, α_2, β_2, $A_2, \cdots, I_2$. If $b_2(t)$ is defined by an equation similar to (21), then

$$p_2 = \bar{p}_2(t) + K_{21}e^{m_{21}t} + K_{22}e^{m_{22}t} + K_{23}e^{m_{23}t} + K_{24}e^{m_{24}t}$$

in which the K's and m's have meanings analogous to those in equation (20). As soon as p_1 and p_2 are known we have u_1 and u_2 from the demand-supply equation (15).

The differential equation which gives $p_0(t)$ is $dG_{p'}/dt = G_p$, which in terms of the coefficients of the cost function becomes

$$-2E_0p'' + 2\delta E_0p' + (2H_0 - 2d_0 - k_0\delta)p + \delta G_0 + I_0 = e_0.$$

If m_{01} and m_{02} are the roots of $2E_0m^2 - 2\delta E_0m - 2H_0 + 2d_0 + k_0\delta = 0$, the solution for p_0 may be written

$$p_0 = \bar{p}_0(t) + K_{01}e^{m_{01}t} + K_{02}e^{m_{02}t},$$

where $\bar{p}_0(t)$ is any solution of the differential equation and K_{01}, K_{02} are arbitrary constants. In particular, if e_0 is a constant, this solution may be taken $\bar{p}_0 = (e_0 - G_0\delta - I_0)/2H_0$. However, the finding of a particular solution does not depend on e_0 being a constant since there are many functions of t which put in the place of e_0 would yield a particular solution easily.

7. *Determining the end values.* We now use the conditions on the end values t_1, $u_1(t_1)$, $p_1(t_1)$, w_1, $u_1(w_1)$, $p_1(w_1)$, $p_0(w_1)$, w_2, $p_0(w_2)$, $u_2(w_2)$, $p_2(w_2)$, t_2, $u_2(t_2)$, $p_2(t_2)$, to determine the constants t_σ, w_σ, $K_{\sigma1}$, $K_{\sigma2}$, $K_{\sigma3}$, $K_{\sigma4}$, K_{01}, K_{02}. These end values are subject to the transversality conditions (14). Since $w_1 = x_2 = x_3$, $x_4 = x_5 = w_2$, we must add the equations $\psi_6 = x_2 - x_3 = 0$, $\psi_7 = x_4 - x_5 = 0$, to the known end conditions (16) in evaluating the matrix (14). We now find that every determinant of order 8 of the $(8 + 18)$-dimensional matrix

$$\left\| \begin{array}{ccccccccc} N_1 & -F_{u_1'}(t_1) & -F_{p_1'}(t_1) & N_2 & N_3 & N_4 & N_5 & -H_{u_2'}(w_2) & c_k \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 \end{array} \right\|$$

$(k=1,\cdots,10)$, must vanish. In this matrix $c_1 = F_{u_1'}(w_1)$, $c_2 = F_{p_1'}(w_1)$, $c_3 = -G_{u_0'}(w_1) \equiv 0$, $c_4 = -G_{p_0'}(w_1)$, $c_5 = G_{u_0'}(w_2) \equiv 0$, $c_6 = G_{p_0'}(w_2)$, $c_7 = -H_{p_2'}(w_2)$, $c_8 = N_6 = -H_{u_2'}(t_2)u_2'(t_2) - H_{p_2'}(t_2)p_2'(t_2) + H(t_2)$, $c_9 = H_{u_2'}(t_2)$, $c_{10} = H_{p_2'}(t_2)$. If it is assumed that the time T to which all profits are discounted is t_1, necessary and sufficient conditions that every determinant of order 8 vanish are

$$N_2(w_1,w_2) + N_3(w_1) + N_4(w_2) + N_5(w_1,w_2) = 0, \quad c_k = 0,$$

$(k=1,\cdots,10)$, which are equivalent to the following equations:

$$+N_3 + N_4 + N_5 + c_1u_1'(w_1) + c_2p_1'(w_1) + c_4p_0'(w_1) + c_6p_0'(w_2) + c_7p_2'(w_2)$$
$$= F(w_1) - G(w_1) + G(w_2) - H(w_2) + H_{u_2'}(w_2)u_2'(w_2) + \int_{w_1}^{w_2} (G_{w_1} + G_{w_2})\,dt = 0,$$

$$(22) \qquad \begin{aligned}
&c_1 = [-2D_1u_1'(w_1) - F_1]\,E(t_1,w_1) = 0,\\
&c_2 = [-2E_1p_1'(w_1) - G_1 - h_1\lambda_{11}(w_1)]\,E(t_1,w_1) = 0,\\
&c_3 \equiv c_5 \equiv 0,\\
&c_4 = [2E_0p_0'(w_1) + G_0]\,E(t_1,w_1) = 0,\\
&c_6 = [-2E_0p_0'(w_2) - G_0]\,E(t_1,w_2) = 0,\\
&c_7 = [2E_2p_2'(w_2) + G_2 + h_2\lambda_{12}(w_2)]\,E(t_1,w_2) = 0,\\
&c_8 + c_9u_2'(t_2) + c_{10}p_2'(t_2) = H(t_2) = 0,\\
&c_9 = [-2D_2u_2'(t_2) - F_2]\,E(t_1,t_2) = 0,\\
&c_{10} = [-2E_2p_2'(t_2) - G_2 - h_2\lambda_{12}(t_2)]\,E(t_1,t_2) = 0.
\end{aligned}$$

Since u and p are expressible in exponentials in t, the integration indicated in the first of these equations can be performed without difficulty and the explicit expression in terms of the given constants can then be exhibited as has been done in the other equations.

The fourteen constants t_σ, w_σ, $K_{\sigma 1}$, $K_{\sigma 2}$, $K_{\sigma 3}$, $K_{\sigma 4}$, K_{01}, K_{02}, can now be determined by the five equations (16) and the nine equations (22), the system (22) giving us only nine equations since c_3 and c_5 are identically zero.

Interesting interpretations can be given some of the end-conditions. Since $H(t)$ represents the profits per unit time from the replacement machine, the condition $H(t_2) = 0$ means that the replacement machine should be run until the amount of money received for the goods sold equals the cost of production at that time. From the condition $u_1'(w_1) = -F_1/2D_1$ the slope of the production curve at the time the operating machine is scrapped is seen to be a constant which depends only on the coefficients of the cost function of the machine. Roos has shown that in typical cases we have $D_1 > 0$ and $F_1 \geqq 0$.*

---

* See Roos, "Some Problems of Business Forecasting," *Proceedings of the National Academy of Sciences*, Vol. 15 (March, 1929), p. 190.

Hence the rate of production is decreasing at this time. Similar conditions on the rate of production at the time of scrapping the replacement machine follow from the equation $u_2'(t_2) = -F_2/2D_2$.

8. *Other forms of the problem.* Evans has suggested that in certain cases the demand depends partly on the seasons, and he has given a form of the demand-supply equation which involves a periodic term as follows: $y = ap + b + b' \cos kt + h dp/dt$, $b' < b$, a, b, b', h all constants.* It will be noticed that if in equation (21), $\mu_{11} = 0$, and μ_{12}, μ_{13} are pure imaginaries (hence, equal except for sign, since we assume all coefficients to be real) the resulting demand-supply equation is in Evans' form.

There is also a variety of other forms of this function $b(t)$ in the demand-supply equation which give a readily integrable differential equation (19). In particular, if b is any exponential in the first power of t, or a polynomial in t, or a constant is this true.

The end equations (16) could also be replaced by other conditions without altering the analysis of the problem. It will be noticed that the conditions (22) can be simplified by assuming that more of the end values are known constants.

* G. C. Evans, "The Mathematical Theory of Economics," *American Mathematical Monthly*, Vol. 32 (1925), p. 108.

A Prepared System for Two Quinary Quadratic Forms.

By J. Williamson.

Introduction. In a previous paper,* a prepared system was determined, in terms of which every concomitant of two quadratics in n variables could be expressed, if the concomitants were multiplied by suitable invariant factors. In this paper we determine a prepared system, for the case $n = 5$, in terms of which every concomitant can be expressed, without being multiplied by an invariant factor. We find that eight new factors must be added to the $2^5 - 1 = 31$ factors already determined, giving a total of 39. In addition a complete list of several types of irreducible concomitants is obtained.

We use the notation of the previous paper throughout except that, for convenience in printing, primes are now used to denote determinantal permutations; i. e. $(ab'c')d_x'$ is used instead of $(a\dot{b}\dot{c})\dot{d}_x$ to denote the series $(abc)d_x - (abd)c_x - (adc)b_x$. Furthermore, for the five sets of cogredient point variables, that are necessary for this discussion, we use x, y, z $t, w,$ and write P, p, and u for the compound coördinates π_2, π_3, and π_4 respectively.†

In the first two sections the results are listed, while the remaining sections are devoted to their determination.

1. *The Prepared System.* This system consists of 5 x-factors, 5 u-factors, 10 P-factors, 10 p-factors, the factor (12345), 3 px-factors, 3 Pu-factors and 2 xu-factors. A complete list of these factors is given below.

$1_x = a_x,\quad 5_x,\quad 2_x = (A\rho x) = a_\rho' b_x',\quad 4_x,\quad 3_x = (A_3R_3x) = (a'b'R_3)c_x'$
$(12) = a_\rho(AP),\quad (54),\quad (13) = (aA_3R_3P) = (a'b'R_3)(ac'P),\quad (53),$
$(14) = (aR\alpha P) = r_\alpha'(as'P),\quad (52),\quad (15) = (arP),$
$(23) = (AR_3)(A_3\rho P) = (AR_3)a_\rho'(b'c'P),\quad (43),$
$(24) = (A\rho R\alpha P) = a_\rho'\overline{r}_\alpha(b'\overline{s}P);$
$(123) = a_\rho(AR_3)(A_3p),\quad (543),\quad (124) = a_\rho(AR\alpha p) = a_\rho r_\alpha'(As'p),$
$(542),\quad (125) = a_\rho(Arp),\quad (541),\quad (143) = (RA_3)(aR_3\alpha p) = (RA_3)r_\alpha'(as't'p),$
$(523),\quad (135) = (aA_3R_3rp) = (a'b'R_3)(ac'rp),$
$(234) = (AR_3)(RA_3)(\alpha\rho p) = (AR_3)(RA_3)a_\rho'(b'c'd'p);$

---

* J. Williamson, "A Special Prepared System for Two Quadratics in n variables," *American Journal of Mathematics*, Vol. 52 (April, 1930), pp. 399-412.

† *Loc. cit.*, §§ 1 and 2.

$(1234) = a_\rho (AR_3)(RA_3)u_\alpha$, (5432), $(1235) = a_\rho (AR_3)(A_3ru)$,
(5431), $(1254) = a_\rho r_\alpha (ARu)$; $(12345) = a_\rho r_\alpha (AR_3)(RA_3)$;
$(12, 54) = 1_x'(2'54) = (125')4_x' = a_\rho r_\alpha (ARpx) a_\rho r_\alpha a_x'(b'Rp) = a_\rho r_\alpha s_x'(Ar'p)$,
$(12, 43) = 1_x'(2'43) = (124')3_x' = a_\rho (RA_3)(AR_3\alpha px) = a_\rho (RA_3) r_x' s_\alpha'(At'p)$,
$(54, 23) = 5_x'(4'23) = (542')3_x'$;
$(123, 543) = (2'3)(1'543) = (4'3)(1235') = a_\rho (AR_3) r_\alpha (RA_3)(A_3R_3Pu)$,
$\qquad = a_\rho (AR_3) r_\alpha (RA_3)(b'c'P)(a'R_3u) = a_\rho (AR_3) r_\alpha (RA_3)(r's'P)(A_3t'u)$
$(123, 154) = (12')(1543') = (14')(1235') = a_\rho (AR_3) r_\alpha (A_3aRPu)$,
$\qquad = a_\rho (AR_3) r_\alpha (a'b'P)(aRc'u) = a_\rho (AR_3) r_\alpha (as'P)(A_3r'u)$,
$(543, 512) = (54')(5123') = (52')(5431')$;
$(12, 543) = 1_x'(5432') = 3_x'(5'4'12) = a_\rho r_\alpha (RA_3)(R_3Aux)$,
$\qquad = a_\rho r_\alpha (RA_3) a_x'(R_3b'u) = a_\rho r_\alpha (RA_3) t_x'(r's'Au)$,
$(54, 123) = 5_x'(1234') = 3_x'(1'2'54)$.

In the above list A, R, α, ρ have been written for A_2, R_2, A_4, R_4 respectively and $A = ab$, $A_3 = abc$, $R = rs$, $R_3 = rst$. When two factors are similar, only one has been defined, since the other may be obtained by replacing a, A, A_3, α by r, R, R_3, ρ respectively.

2. *Complete list of irreducible concomitants of several types.*

6 invariants: $(a\alpha)^2$, $(a\rho)^2$, $(AR_3)^2$, $(RA_3)^2$, $(r\alpha)^2$, $(r\rho)^2$.

6 covariants: 5 quadratics i_x^2 and 1 quintic $(12345)1_x2_x3_x4_x5_x$.

6 contravariants: 5 quadratics $(ijkm)^2$,
1 quintic $(1234)(1235)(1245)(1345)(2345)$.

20 complexes containing the variable P:
10 quadratics $(ij)^2$, 10 cubics $(ij)(jk)(ki)$.

20 complexes containing the variable p:
10 quadratics $(ijk)^2$, 10 cubics $(12345)(ijk)(ijm)(ijn)$.

44 mixed forms containing u and x:
5 of orders 1 in u and 1 in x, $(12345)(1234)5_xS$,
$(12345)(1245)3_x$, $(12345)(54, 123)S$.
5 of orders 1 in u and 4 in x, $(ijkm)i_xj_xk_xm_x$.
5 of orders 4 in u and 1 in x,
$(12345)(mijk)(mijn)(mikn)(mjkn)m_x$.
10 of orders 2 in u and 3 in x, $(12345)(ijkm)(ijkn)i_xj_xk_x$.
10 of orders 3 in u and 2 in x, $(mnkj)(mnij)(mnik)m_xn_x$.
2 of orders 3 in u and 3 in x, $(12345)(1245)(1345)(2345)1_x2_x3_xS$,
2 of orders 3 in u and 4 in x, $(12, 543)(1245)(1345)1_x4_x5_xS$.

4 of orders 4 in u and 4 in x,
$(12, 543)(12345)(1234)(1245)(1345)1_x4_xS$,
$(12, 543)(12345)(1235)(1245)(1345)1_x5_xS$,
1 of orders 5 in u and 4 in x,
$(12, 543)(54, 123)(1235)(1245)(1345)1_x5_x$.

67 mixed forms containing P and x:
10 of orders 1 in P and 2 in x, $(ij)i_xj_x$,
4 of orders 1 in P and 3 in x,
$(12345)(12)3_x4_x5_xS$, $(12345)(23)1_x5_x4_xS$,
5 of orders 2 in P and 1 in x,
$(12345)(23)(45)1_xS$, $(12345)(43)(15)2_xS$, $(12345)(12)(45)3_x$,
5 of orders 2 in P and 3 in x,
$(12345)(21)(15)1_x3_x4_xS$, $(12345)(12)(23)2_x4_x5_xS$,
$(12345)(23)(34)3_x1_x5_x$,
20 of orders 3 in P and 1 in x,
$(12345)(12)(14)(15)3_xS$, $(12345)(21)(23)(25)4_xS$,
$(12345)(31)(32)(34)5_x$, $(12345)(12)(13)(45)1_xS$,
$(12345)(12)(15)(43)1_xS$, $(12345)(14)(15)(23)1_xS$,
$(12345)(21)(23)(54)2_xS$, $(12345)(24)(23)(51)2_xS$,
$(12345)(25)(21)(43)2_xS$, $(12345)(34)(35)(21)3_xS$,
$(12345)(34)(32)(15)3_x$.
2 of orders 3 in P and 3 in x, $(12345)(34)(32)(45)1_x3_x4_xS$.
18 of orders 4 in P and 3 in x,
$(12345)(ij)(ik)(im)(in)i_x$ 5 in number,
$(12345)(13)(15)(32)(34)1_xS$, $(12345)(14)(12)(43)(45)1_xS$,
$(12345)(15)(12)(53)(54)1_x$, $(12345)(21)(23)(14)(15)2_xS$,
$(12345)(24)(21)(43)(45)2_xS$, $(12345)(31)(34)(12)(15)3_xS$,
$(12345)(32)(34)(21)(25)3_xS$.
1 of orders 4 in P and 3 in x, $(12345)(12)(23)(34)(45)2_x3_x4_x$.
2 of orders 5 in P and 1 in x, $(12345)(31)(14)(23)(34)(45)1_xS$.

67 mixed forms containing p and u: These forms are the duals * of the mixed forms containing P and x and can be written down immediately. For example from the 5 forms

$$(12345)(ij)(ik)(im)(in)i_x,$$

we obtain the 5 dual forms

$$(kmn)(jmn)(jkn)(jkm)(jkmn).$$

* *Loc. cit.*, § 5.

In the above list i, j, k, m, n take the values 1, 2, 3, 4, 5 with the understanding that in any one form i, j, k, m, n are all distinct. The presence of the letter S after a form denotes the existence of a similar form,* that is a form in which the symbols 1 and 2 are interchanged with the symbols 5 and 4 respectively. To obtain the actual irreducible concomitants from this list we must remove from any form the invariant factors which appear. For example, $(12)^2 = a_\rho^2(AP)^2$ yields the actual concomitant $(AP)^2$.

3. *Determination of the Prepared System.* Since we are now considering two quadratics in n variables for the case $n = 5$, there are six invariants † and five quadratic covariants † i_x^2, $(i = 1, 2, 3, 4, 5)$. By theorem I every concomitant,‡ multiplied by a suitable invariant factor, can be expressed in terms of the symbolic factors,

$$i_x, \quad (ij), \quad (ijk), \quad (ijkm), \quad (12345) \qquad (i, j, k, m = 1, 2, 3, 4, 5).$$

We must now determine, if ever in forming these bracket factors, we have disturbed any of the invariant factors, which appear when 12, 23, 34, or 45 are convolved together. Originally we have five sets of cogredient point variables x, y, z, t, w, which are convolved as $\Delta = (xyztw)$, $u = xyzt$, $p = xyz$, $P = xy$. Since the only factor involving Δ is (12345) and since 12, 23, 34, 45, are all convolved in this, no invariant factor has been disturbed in forming it. When all the variables w have been convolved with $xyzt$ to form Δ, we are left to consider 4-factors, 3-factors and 2-factors, where an i-factor is a factor involving i of the variables x, y, z, t. We may neglect all 4-factors, as they lead to nothing new, for then the variables can only be $xyzt = u$. Let us now consider the possible cases, in which x, y, z, t may be convolved to form u. If one of these variables occur in a 3-factor, we may assume that three of them occur in this 3-factor, for $(ijk \mid x'zt)(rs \mid y'\xi) \equiv (ijkr' \mid xyzt)(s' \mid \xi) +$ terms in which xy are convolved together, and $(ijk \mid xy'z')(rs \mid t'\xi) \equiv (ijkr' \mid xyzt)(s' \mid \xi) +$ terms in which yzt are convolved together. Hence we must consider the cases when three variables occur in a 3-factor and the fourth occurs (a) in a 3-factor and (b) in a 2-factor.

Case (a) gives the possibility, $(ijkn' \mid u)(im' \mid \xi\eta)$, and case (b) $(ijkn' \mid u)(m' \mid \xi)$, where ξ, η may be any of x, y, z, t. In (a) and (b), neither of m, n is the same as any of i, j, k or else no convolution of successive integers is disturbed. At first sight it would appear that $(ijk'\dot{r}|u)(m'|\xi)(\dot{n}|\eta)$

* *Loc. cit.*, § 6.

† *Loc. cit.*, p. 404.

‡ *Loc. cit.*, § 3.

is a possibility, arising from three 2-factors, but i, j, k, r, m, n must all be distinct, and this is impossible. But if the variable t does not appear, we might have the single new type (c) $(ijk' \mid p)(m' \mid \xi)$, arising from two 2-factors.

We now write the factors for simplicity without the variables, since no confusion can arise. There are no further types of factors, as we shall see. Type (a) cannot occur with another u-factor, as

$$(ijkn')(im'rs), \quad (ijkn')(im'\dot{r}\dot{s})\dot{t}$$

are the only possibilities. In the first rs cannot contain i, m or n, and so must be jk. But by the fundamental identities this is impossible.* In the second case none of r, s, t can be i, therefore two of them must be either k, j or m, n and in either case no invariant factor is disturbed. Further since *

$$(ijkn')(irm') \equiv (ijkr)(imn) + (ijk)(inmr),$$

type (a) cannot occur with a further p-factor. Hence type (a) gives solely the one new factor type $(ijkn')(im')$.

Similarily it may be shown that type (b) cannot occur with another u or p-factor. Moreover type (c) cannot occur with another p-factor, for *

$$(ijk')(n'm) \equiv (ijm)(nk) + (ij)(knm),$$

and in both terms on the right i, j and n, k are convolved. There are three other possible cases to consider; $(ijkn')(m'a)$ from (b) and a 1-factor, $(ijkn')(m'\dot{a})(\dot{b}cde)$ from two (b) factors, and $(ijk')(m'ab)$ from one (c) factor and a 2-factor. Of these, the first reduces to type (a), since $a =$ one of i, j, k; the third gives nothing new, since one of a, b must be i or j; the second is more easily treated later.

In type (a), m, n must be consecutive integers and so must i, j. Hence we have the possibilities;

$$(1235')(14') \equiv (123, 154), \qquad (3215')(3'4') \equiv (123, 543),$$
$$(5431')(52') \equiv (543, 512).$$

In type (b) m, n must be consecutive integers and so we have;

$$1_x'(2'345), \quad 2_x'(3'145), \quad 3_x'(4'125), \quad 5_x'(4'321).$$

But

$$2_x'(3'145) \equiv 1_x(3245) + (3214')5_x' \equiv 5_x'(3214'),$$

since in (3245) both 2, 3 and 4, 5 are convolved. Similarly

$$3_x'(4'125) \equiv 1_x'(432'5).$$

---

* *Loc. cit.* Formulas (16) and (17).

Accordingly type (b) yields only two new factors,

$$1_x'(5432') \equiv (12, 543), \quad 5_x'(1234') \equiv (54, 123).$$

In type (c) both i, j and m, k must be successive integers and so we have,

$$1_x'(2'43) \equiv (12, 43), \quad 5_x'(4'23) \equiv (54, 23), \quad 1_x'(2'54) \equiv (12, 54).$$

If a new type of factor arises from two (b) factors, it must be

$$(1'\dot{4})(2'345)(\dot{5}321) \equiv (123'\dot{4})(4'5')(\dot{5}321),$$
$$\equiv (1254)(4'5')(3'321) \equiv (1254)(34')(5'321),$$

and so is expressible in terms of simpler factors. We thus have the eight new factors, three of type (a), three of type (c), and two of type (b). The factors of type (c) are the duals * of those of type (a), while each of the factors of type (b) is the dual of the other.

Now, since, with the addition of these new factor types, no invariant factors, which were originally introduced,† have been disturbed, we can work with the symbols i, j etc. and at the end remove all actual invariant factors and obtain the actual irreducible concomitants, provided that no identity is used, which separates successive integers convolved an even number of times. An alternative method is to use as a prepared system the factors (AP) for (12) etc. This prepared system was actually found by Dr. Wm. Saddler, but has never been published. He determined the prepared system by methods analogous to those used by H. W. Turnbull in his paper on two quadratics in four variables.‡ To find any of the irreducible concomitants by this method is cumbersome, as all identities have to be worked out in detail and in addition the ten symbols a, r, A, R, A_3, R_3, α, ρ must be paired off instead of the five symbols 1, 2, 3, 4, 5. This, together with the simplification of the identities, more than compensates for the addition of the extra factor (12345) and the fact that the identities cannot be applied blindly.

4. *Determination of the irreducible covariants and contravariants.* The factors which may occur in a covariant are the five i_x factors and (12345). The irreducible covariants are then six in number, the five quadratics i_x^2, and the quintic $(12345)1_x2_x3_x4_x5_x$. By duality § the contravariants are also six in number, the five quadratics $(kmnj)^2$ and the quintic

* *Loc. cit.*, § 5.

† *Loc. cit.*, p. 405.

‡ H. W. Turnbull, "The Simultaneous System of Two Quadratic Quaternary Forms," *Proceedings of the London Mathematical Society*, Ser. 2, Vol. 18, Parts 1 and 2, pp. 70-94.

§ *Loc. cit.*, § 5.

$$(1234)(1235)(1245)(1345)(2345).$$

5. *Determination of the irreducible complexes.* The possible factors, which may occur, are the ten factors (ij) and (12345). But, as a product of (12345) by factors of the type (ij) always involves an odd number of symbols, the factor (12345) cannot appear in such a concomitant. Since the factors (ij) are strictly analogous to simple bracket factors of binary forms, we have only 20 possible complexes,

$$\text{the 10 quadratics } (ij)^2 \text{ and the 10 cubics } (ij)(jk)(ki),$$

for a product of four or more factors (ij) is reducible. In fact,

$$(ij)(km) \equiv (ik)(jm) + (kj)(im), \quad (ij)(kn) \equiv (ik)(jn) + (kj)(in).$$

By multiplying these two equations together and neglecting the terms, which involve a factor squared, we have

$$(ni)(ik)(kj)(jm) + (mi)(ik)(kj)(jn) \equiv 0,$$

or

$$2(ni)(ik)(kj)(jm) + (ji)(ik)(kj)(mn) \equiv 0,$$

by applying the identity $(mi')(j'n') \equiv 0$ to the second term. But as $(ji)(ik)(kj)$ is itself a concomitant $(ni)(ik)(kj)(jm)$ is reducible.*

By the principle of duality we see that there are only 20 irreducible complexes involving the variable p, the 10 quadratics $(ijk)^2$ and the 10 cubics $(12345)(ijk)(ijm)(ijn)$.

6. *Determination of the mixed concomitants containing u and x. Reductions.* (a) Since $(12, 543) \equiv 1_x'(5432') \equiv 3_x'(5'4'12)$, any concomitant containing the factor $(12, 543)$ is reducible, if 12 or both of 34, 45 are convolved an odd number of times. In addition such a concomitant has a factor $(AR_3)(AR_3ux)$ if 23 is convolved an odd number of times.

Further,

$$(12, 543) \equiv 3_x'(5'4'12) \equiv 3_x(5412) - 5_x'(34'12) \equiv 3_x(5412) - (54, 123).$$

Hence

(b) $$(12, 543)(1254)3_x \equiv (12, 543)(54, 123) \equiv 0, \text{ by (a)}.$$

There also exists a reduction similar to that for quaternary forms.†

(c) $$(12, 543)(5432)2_x \equiv 0.$$

---

* Grace and Young, *Algebra of Invariants*, Chap. 15, p. 322.

† J. Williamson, "Note on the Simultaneous System of Two Quadratic Quaternary Forms, *Journal of the London Mathematical Society*, Vol. 4 (1929), pp. 182-183.

For neglecting the invariant factors we have

$$(12, 543)(53432)2_x \equiv (AR_3ux)(A\rho x)u_\rho = 2(aR_3u)b_x[a_\rho b_x - a_x b_\rho]u_\rho,$$

and each term on the right has a factor b_x^2 or $b_\rho u_\rho b_x$. It is important to notice that the dual product $(54, 123)1_x(1345)$ is not reducible. Moreover

(d) $(1234)5_xM \equiv 0$, if 4, 5 is convolved an odd number of times in M, and $(2345)1_xM \equiv 0$, if 1, 2 is convolved an odd number of times in M.

Since $(12, 543) \equiv (5432)1_x - (5431)2_x$, by squaring this identity

(e) $$(5432)(5431)1_x2_x \equiv 0.$$

If now we consider (5432) as simpler than (5431),

(f) $(5431)2_xM \equiv 0$, if 2, 3 is convolved an odd number of times in M.

Again by squaring the identity

$$(12, 543) - (2431)5_x \equiv (5231)4_x + (5421)3_x,$$

we have,

(g) $$(5231)(5421)4_x3_x \equiv 0 \text{ by (a)}.$$

In the above reductions we may replace each factor by its similar factor and in most cases obtain a new reduction. We now consider the possible forms in the following order; first those without the factor $(12, 543)$ and in ascending order in u, then those with one factor $(12, 543)$ and finally those with both factors $(12, 543)$ and $(54, 123)$. We only write down one of each pair of similar forms, and those forms which are marked R are reducible. The method of reduction is indicated shortly at the side.

One u factor. We have the five concomitants

$(ijkm)i_xj_xk_xm_x$,

and the types

$(12345)(1234)5_x$, $(12345)(1235)4_x\,R$(f), $(12345)(1245)3_x$.

Two u factors. We have the types

$(1234)(1235)4_x5_x\,R$(e), $(1234)(1245)3_x5_x\,R$(d),

$(1234)(2345)1_x5_x\,R$(d), $(1234)(1345)2_x5_x\,R$(d) and (f),

$(1235)(1245)3_x4_x\,R$(g), $(1235)(1345)2_x4_x\,R$(f) and (a),

and the ten $(12345)(ijkm)(ijkn)i_xj_xk_x$.

Three u factors. We have the ten $(mnjk)(mnij)(mnik)m_xn_x$, the duals of the previous case and the types

$(12345)(1245)(1345)(2345)1_x 2_x 3_x,$
$(12345)(1234)(1235)(1345)2_x 4_x 5_x \quad R(\mathrm{f}),$
$(12345)(1423)(1425)(1435)2_x 3_x 5_x \quad R(\mathrm{d}),$
$(12345)(1523)(1524)(1534)2_x 3_x 4_x \quad R$ by $(1524)3_x,$
$(12345)(2314)(2315)(2345)1_x 4_x 5_x \quad R(\mathrm{d}),$
$(12345)(2415)(2413)(2435)1_x 3_x 5_x \quad R(\mathrm{d}).$

Four u factors. We have the types

$(1234)(2315)(1245)(1345)2_x 3_x 4_x 5_x \quad R(\mathrm{e}),$
$(1234)(1235)(1245)(2345)1_x 3_x 4_x 5_x \quad R(\mathrm{d}),$
$(1234)(1253)(1345)(2345)1_x 2_x 4_x 5_x \quad R(\mathrm{f}),$

and the five $(12345)(ijkm)(ijkn)(ijmn)(ikmn)i_x$, the duals of the forms with one u and four x factors.

Five u factors. Either no x factors or five x factors occur. The first case has already been considered and in the second case all possible forms are reducible.

One factor $(12, 543)$. We have the simple forms

$(12345)(12, 543)$, $(12, 543)^2$ and two similar forms.

One further u factor. By (a) we see that there is only one possibility $(12, 543)(1245)3_x \quad R(\mathrm{b})$.

Two further u factors. We have the types,

$(12, 543)(1234)(2345)2_x 3_x 4_x \quad R(\mathrm{c}),$
$(12, 543)(1235)(2345)5_x 3_x 2_x \quad R(\mathrm{c}),$
$(12, 543)(1245)(1345)1_x 4_x 5_x,$
$(12, 543)(12345)(1245)(1234)3_x 5_x \quad R(\mathrm{d}),$
$(12, 543)(12345)(1245)(1235)3_x 4_x \quad R(\mathrm{f})$ mod $(12, 543)(54, 123),$
$(12, 543)(12345)(2345)(1345)1_x 2_x \quad R(\mathrm{f})$ mod $(12, 543)^2.$

Three further u factors. We have the types

$(12, 543)(1234)(1235)(1245)3_x 4_x 5_x \quad R$ mod $(12, 543)(54, 123),$
$(12, 543)(1234)(2345)(1345)1_x 2_x 5_x \quad R$ mod $(12, 543)^2,$
$(12, 543)(12345)(1234)(1235)(2345)2_x 3_x \quad R(\mathrm{c}),$
$(12, 543)(12345)(1234)(1245)(1345)1_x 4_x,$
$(12, 543)(12345)(1235)(1245)(1345)1_x 5_x.$

Four further u factors. We have the sole possibility

$$(12, 543)(12345)(1345)(2345)(1234)(1235)1_x 2_x 4_x 5_x \qquad R \text{ mod } (12, 543)^2.$$

Five further u factors. There are no irreducible forms of this type. There are thus six irreducible forms containing one factor of the type $(12, 543)$, the

three in the list above and three similar forms. It is important to notice that each of these six is irreducible but that their duals reduce by (c).

Both factors (12, 543) *and* (54, 123). If both the factors (12, 543) and (54, 123) appear in a concomitant, since 12, 54, 23, and 34 must all be convolved an even number of times, there are very few possibilities and finally we are left with

$$(12, 543)(54, 123)(1235)(1245)(1345)1_x 5_x,$$

and its dual

$$(12, 543)(54, 123)(12345)(1234)(2345)2_x 3_x 4_x \quad R(c).$$

7. *Determination of the mixed concomitants containing P and x. Reductions.* If we consider the P factors in the order of simplicity,

$$(12), (54), (23), (34), (15), (14), (52), (13), (53), (24),$$

by identities of the type $(i'j')k_x' \equiv 0$, we see that

(h)

the products $(13)2_x$, $(53)2_x$, $(53)4_x$, $(13)4_x$, $(24)3_x$,
$(24)1_x$, $(25)1_x$, $(42)5_x$, $(41)5_x$,

are reducible. Further by identities of the type $(ij')(k'm') \equiv 0$, we see that the products

(i)

$(13)(24)$, $(53)(24)$, $(13)(52)$, $(53)(14)$, $(14)(25)(15)$,
$(24)(14)(15)$, $(42)(52)(51)$, $(14)(24)(25)$

are reducible. The concomitant

(j) $$(13)(35)(51)M \equiv 0$$

also, since $(13)(35)(51)$ is an actual concomitant containing no invariant factors.

We first consider those concomitants, which do not contain the factor (12345).

No factor (12345). In this case the only irreducible concomitants that appear are the ten mixed forms $(ij)i_x j_x$. This follows as a result of the analogy with binary forms (See § 5).

Forms containing the factor (12345). If the factor (12345) appears in a concomitant M, M must contain five x factors, three x factors or one x factor, since the number of symbols in a P factor is even. Five x factors cannot occur in M, for in that case the symbols appearing in the P factors must be paired

off. Accordingly M contains at least one factor (ij), in which i, j are not successive integers, and as a result the concomitant factor $(ij)i_x j_x$.

By considering list (i), we see that there can be no irreducible concomitants involving eight or more P factors. But if seven P factors occur and (24) occurs, (13) and (53) cannot appear nor can any of the products,

$$(14)(25)(15),\ (24)(15)(14),\ (24)(51)(52),\ (24)(14)(25)$$

and so in this case seven P factors cannot occur. But, if (24) does not occur, the products

$$(13)(52),\ (53)(14),\ (13)(35)(52),\ (14)(25)(15),\ (13)(35)(51)$$

are prohibited. Hence the only possible form involving seven P factors is $(12)(23)(34)(45)(25)(15)(35)M$ or the similar form. Since 1_x must occur among the x factors in M, this form *is* reducible by (h). Hence there are no irreducible concomitants containing seven P factors.

We shall now consider the remaining concomitants in ascending order in P. Since (12), (54); (13), (53); (23), (43); (25), (41); are similar factors and (15), (24) self similar factors, we need only write down one of every two similar forms.

One P factor. There is only one type, $(12345)(ij)k_x m_x n_x$ but all concomitants of this type are equivalent to

$$(12345)(12)3_x 4_x 5_x,\quad (12345)(23)1_x 4_x 5_x$$

and two similar forms by reductions (h).

Two P factors, one x factor. There is only one type $(12345)(ij)(km)n_x$. If we let $n = 1, 2, 3$ in succession and use reductions (h), we are left with

$$(12345)(23)(45)1_x,\ (12345)(43)(15)2_x,\ (12345)(12)(45)3_x,$$

and two similar forms.

Two P factors, three x factors. There is only one type $(12345)(ij)(ik)m_x n_x i_x$. Since, if i, k are not successive integers,

$$(12345)(ij)(ik)m_x n_x i_x \equiv (12345)(ij)(im)k_x n_x i_x \text{ etc.},$$

by letting $i = 1, 2, 3$ in turn we see that all forms of this type reduce to

$$(12345)(12)(15)1_x 3_x 4_x,\ (12345)(21)(23)2_x 4_x 5_x,\ (12345)(32)(34)3_x 1_x 5_x,$$

and two similar forms.

Three P factors, one x factor. There are two possible types

$$(1234)(ij)(ik)(im)n_x,\ (12345)(ij)(ik)(mn)i_x.$$

The concomitants of the first type reduce by reductions (h) to

$$(12345)(12)(14)(15)3_x,\ (12345)(21)(23)(25)4_x,\ (12345)(31)(32)(34)5_x,$$

and two similar forms. For, since the form

$$\begin{aligned} X &= (12345)(31)(32)(34)5_x, \\ &\equiv (12345)(35)(34)(32)1_x + (12345)(15)(32)(34)3_x = Y + Z, \end{aligned}$$

and the form Z appears in the list of irreducible forms of the second type, we may neglect Y, the form similar to X. The concomitants of the second type are equivalent to

$$\begin{array}{ll} (12345)(12)(13)(45)1_x, & (12345)(12)(15)(43)1_x, \\ (12345)(14)(15)(23)1_x, & (12345)(21)(23)(54)2_x, \\ (12345)(24)(23)(51)2_x, & (12345)(25)(23)(14)2_x \quad R, \\ (12345)(25)(21)(43)2_x, & (12345)(34)(35)(21)3_x, \\ (12345)(34)(32)(15)3_x, & \end{array}$$

and seven similar forms. The form marked R reduces by the identity $(25')(1'4') \equiv 0$.

Three P factors, three x factors. There is only one type

$$(12345)(mj)(ji)(ik)i_x j_x n_x.$$

In this i, j; i, k; j, m must all be successive integers, or else the form reduces by the identities $(i'j')n_x' \equiv 0$, $(i'k')j_x' \equiv 0$, $(j'm')i_x' \equiv 0$. Accordingly we are left with $(12345)(34)(32)(45)1_x 3_x 4_x$ and its similar form.

Four P factors, one x factor. There are two types $(12345)(ij)(ik)(im)(in)i_x$, $(12345)(ij)(ik)(jm)(jn)i_x$. The first type yields five concomitants. In the second type, if $i = 1$ and $j = 2$, one of (24) or (25) must occur and so the form is reducible. If $i = 1$ and $j = 3$, (35) cannot occur and so we have the possible form $(12345)(13)(15)(32)(34)1_x$. If $i = 1$ and $j = 4$, (42) cannot occur and if $i = 1$ and $j = 5$, (52) cannot occur and so we have the two forms $(12345)(14)(12)(43)(45)1_x$, $(12345)(15)(12)(43)(45)1_x$, the second of which is equivalent to its similar form. By a similar treatment for the cases $i = 2$, and $i = 3$, we have as a final list of concomitants of the second type

$$\begin{array}{ll} (12345)(13)(15)(32)(34)1_x, & (12345)(14)(12)(43)(45)1_x, \\ (12345)(15)(12)(53)(45)1_x \equiv \text{to its similar form}, & \\ (12345)(21)(23)(41)(15)2_x, & (12345)(24)(21)(43)(45)2_x, \\ (12345)(31)(34)(12)(15)3_x, & (12345)(32)(34)(21)(25)3_x, \end{array}$$

and six similar forms.

Four P factors, three x factors. There are three types

$$(12345)(mn)(ij)(jk)(ki)i_x j_x k_x, \quad (12345)(mi)(ij)(jk)(kn)i_x j_x k_x,$$
$$(12345)(mn)(nj)(jk)(kn)i_x j_x k_x.$$

The first type is not possible, since all of i, j; j, k; k, i cannot be successive integers and so the concomitant resolves into factors. In the second type m, i; i, j; j, k; k, n must be successive integers and so we have only the one irreducible form $(12345)(12)(23)(34)(45)2_x 3_x 4_x$. In the third type n, j; j, k; k, n must all be successive integers and this is impossible.

Five P factors, one x factor. There is only one type

$$(12345)(ij)(ik)(mj)(jk)(kn)i_x.$$

In such a concomitant, by the identity $(j'k')i_x' \equiv 0$, we see that j, k must be successive integers, and by identities of the type $(i'j')(k'n) \equiv 0$, that one factor of each of the products $(ij)(kn)$ and $(mj)(ik)$ must be a pair of successive integers. Since j, k are successive integers, both i, j and m, j cannot be successive integers. If i, j are successive integers, it follows from the above that i, k must be successive integers. But this is impossible since i, j; j, k; k, i cannot all be successive integers. Therefore i, j cannot be successive integers and so k, n must be. Since k, j are also successive integers, i, k cannot be successive integers and so m, j must be. As a result we have only two concomitants of this type $(12345)(13)(14)(23)(34)(45)1_x$ and the similar form.

Five P factors, three x factors. There are four types

$$(12345)(ij)(jk)(kn)(ni)(mn)i_x j_x k_x,$$
$$(12345)(in)(nk)(km)(mi)(mn)i_x j_x k_x,$$
$$(12345)(ij)(jk)(ki)(mi)(in)i_x j_x k_x,$$
$$(12345)(km)(mj)(jk)(mi)(in)i_x j_x k_x.$$

In the first type i, j; j, k; k, n; n, i must all be successive integers, and this is impossible. Similarly the last two types are also impossible. In the second type, if m, n are not successive integers, the form is reducible by the identity $(m'n')i_x' \equiv 0$. Therefore both of k, m and m, i cannot be successive integers. Further either i, n or k, m must be successive integers and also one of the pairs n, k and m, i. If i, n are successive integers, m, i cannot be, since m, n are successive integers, and therefore n, k must be successive integers while k, n cannot be. Hence the form has the factor $(km)(mi)k_x i_x$. If i, n are not successive integers, k, m must be and so m, i cannot be. But, since one of the pairs n, k and m, i must be successive integers, n, k must be suc-

cessive integers. Hence k, m; m, n; n, k must all be successive integers and this is impossible. Accordingly there are no irreducible concomitants containing five P factors and three x factors.

Six P factors and one x factor. There are two types

$$(12345)(jk)(jm)(jn)(km)(kn)(nm)i_x,$$
$$(12345)(ji)(ik)(mj)(jk)(km)(nm)i_x.$$

Of these two types we need only consider the second, for, in the first type one of k, m; k, n; n, m cannot be successive integers and so we can apply an identity of the type $(m'n')i_x' \equiv 0$ and reduce it to two forms of the second type. In the second type we see that k, j must be successive integers and that one factor of each of the products $(mj)(ik)$ and $(mk)(ij)$ must be a pair of successive integers. But it is impossible for this to be the case, since the product of (kj) with one of both pairs (mk), (ij) and (mj), (ik) consist of three factors with a symbol in common or else is of the type $(ij)(jk)(ki)$.

Six P factors and three x factors. There are three types

$$(12345)(im)(mj)(jk)(kn)(ni)(mn)i_xj_xk_x,$$
$$(12345)(ij)(jm)(mk)(ki)(mi)(in)i_xj_xk_x,$$
$$(12345)(im)(mj)(jn)(ni)(mk)(kn)i_xj_xk_x.$$

In the first type m, j; j, k; k, n must be successive integers. Accordingly m, n cannot be successive integers and the form reduces by $(m'n')i_x' \equiv 0$. In the second type m, i; m, j; m, k must all be successive integers and this is impossible. In the third type i, m; m, j; j, n; n, i cannot all be successive integers and so the form reduces to the first type by identities of the type $(i'm')k_x' \equiv 0$. Hence there are no irreducible concomitants containing six P factors. We have already shown that there are no irreducible concomitants with more than six P factors and so the list given in § 2 is complete.

By the principle of duality we can write down the irreducible concomitants involving the variables p and u.

The determination of the concomitants containing the variables P and x has not been attempted. The list of irreducible forms would be considerably longer but could be obtained by the methods that we have used. Since the complete system for two quadratics in four variables contains 122 forms, we should expect to obtain at least 700 or 800 forms in the complete system for two quadratics in five variables, as the labour involved in the latter case is at least five times as heavy as that in the former.

THE JOHNS HOPKINS UNIVERSITY.

Rational Surfaces Defined by Linear Systems of Plane Curves $C_{3n}: 8A^nB^{n-1}$.

By JOSEPH CRAWFORD POLLEY.

1. *Introduction.* The rational surfaces of order four and five having no multiple curves, and those of order five having multiple curves insufficient for rationality, have been determined. There are three types of rational quartic surfaces with no double curve. One was discovered by L. Cremona * by applying a cremona transformation to a known quartic surface. The two remaining types were determined by M. Noether.† His method of investigation was that of considering quartic surfaces with a double point and applying to their equations the conditions for a one to one correspondence with the points of a plane.

Of particular interest also is the work of D. Montesano on rational quintic surfaces.‡ He obtained all the possible types by applying special cremona transformations to known rational surfaces.

In this paper various rational surfaces are discussed by considering certain linear systems of plane curves. The surfaces of Cremona, Noether and some of those of Montesano are re-determined by this method and a general type of surface is discussed by means of a linear system of plane curves of the form $C_{3n}: 8A^nB^{n-1}$.

2. *The system* $C_6: 7A^2$. In a plane (x) the system of curves $C_6: 7A^2$, with double points at 7 points A_i $(i = 1, 2, \cdots, 7)$, is of dimension 6. Let $C_3: 7A$, $C_3': 7A$ and $C_3'': 7A$ be linearly independent members of the net of cubics determined by the points A_i; and $C_6: 7A^2$ a non-composite sextic of the system. Then we can take as the equation of the system

$$(1)\quad a_1C_3^2 + a_2C_3C_3' + a_3C_3C_3'' + a_4C_3'C_3'' + a_5C_3'^2 + a_6C_3''^2 + a_7C_6 = 0.$$

Let

$$(2)\quad \begin{aligned} &y_1 = C_3^2: 7A, \quad y_2 = C_3: 7A \cdot C_3': 7A, \quad y_3 = C_3: 7A \cdot C_3'': 7A, \\ &y_4 = C_3': 7A \cdot C_3'': 7A, \quad y_5 = C_3'^2: 7A, \quad y_6 = C_3''^2: 7A, \quad y_7 = C_6: 7A^2. \end{aligned}$$

* L. Cremona, *Coll. Math. Chelini* 413-424 (1881).

† M. Noether, *Mathematische Annalen*, Vol. 33 (1889), pp. 546-571.

‡ D. Montesano, *Rendiconti della Reale Accademia di Napoli*, Ser. 3, Vol. 13 (1907), pp. 66-68.

These are the parametric equations of a rational surface F_a in S_6, in $(1,1)$ correspondence with the plane (x), the image of a point (x) being a point (y) on the surface. Since any two members of the set have eight residual intersections, a general S_4 in S_6 meets F_a in eight points. Hence F_a is of order eight.

For a general point on C_3

$$(3)\quad y_1 = y_2 = y_3 = 0,\quad y_4 = C_3'C_3'',\quad y_5 = C_3'^2,\quad y_6 = C_3''^2,\quad y_7 = C_6.$$

Hence the image of C_3 is a curve L in a sub-space S_3 of S_6.

Since a $C_6: 7A^2$ has, with $C_3: 7A$, four residual intersections, a plane

$$y_1 = y_2 = y_3 = \beta_4 y_4 + \beta_5 y_5 + \beta_6 y_6 + \beta_7 y_7 = 0$$

meets L in four points. Hence L is a quartic curve of genus 1.

By projection from the plane of $y_1 = y_2 = y_3 = y_7 = 0$ the surface F_8 goes into a surface F_4, in an S_3, whose parametric equations are

$$(4)\quad \begin{array}{ll} y_1 = C_3^2: 7A, & y_2 = C_3: 7A \cdot C_3': 7A, \\ y_3 = C_3: 7A \cdot C_3'': 7A, & y_7 = C_6: 7A^2. \end{array}$$

The image of L is the point $(0,0,0,1)$. To the plane sections of F_4 corresponds the system

$$a_1 C_3^2 + a_2 C_3 C_3' + a_3 C_3 C_3'' + a_4 C_6 = 0$$

all members of which pass through the four simple points in which C_3 meets C_6. The system is therefore of grade 4.

For a point near P_i $(i = 1, 2, 3, 4)$, the residual intersections of C_3 and C_6, the corresponding y_2, y_3 and y_7 are infinitesimals of the first order and the corresponding y_1 is an infinitesimal of the second order. Hence the images of P_i are straight lines in the plane $y_1 = 0$ through the point $(0,0,0,1)$.

The surface (4) is the well known rational quartic surface of Cremona.

3. *The system* $C_7: A^3 8B^2$. In a plane (x) the system of curves $C_7: A^3 8B^2$, with a triple point at A and double points at B_i $(i = 1, 2, \cdots, 8)$, is of dimension 5. The basis points determine a cubic $C_3: A8B$ and a web of quartics $C_4: A^2 8B$. Hence, if $C_7: A^3 8B^2$ and $C_7': A^3 8B^2$ are two non-composite curves of the system, and $C_1: A$, $C_1': A$ are members of the pencil of lines on A, the following can be taken as the linearly independent members of the system:

$$(5)\quad \begin{array}{lll} C_7: A^3 8B^2, & C_3: A8B \cdot C_4: A^2 8B, & C_3^2: A8B \cdot C_1: A, \\ C_3^2: A8B \cdot C_1': A, & C_3: A8B \cdot C_4': A^2 8B, & C_7': A^3 8B^2. \end{array}$$

A member of the pencil $C_7 + \gamma C_7' = 0$ has two residual intersections with C_3 and for a particular choice of γ there is a member tangent to C_3 at some point C. Let C_7' be that member and consider the pencil $C_7' + \delta C_3 C_4'$ $= 0$. For a particular choice of δ we obtain a C_7 with a double point at C; call it $\bar{C}_7$. Furthermore, among the quartics there is a net through the point C, $C_4: A^2 8BC$, $C_3: A8BC \cdot C_1: A$ and $C_3: A8BC \cdot C_1': A$. Let D be the residual intersection of C_4 and C_3. Then we have as linearly independent members of the system

$$(6)\quad \begin{array}{lll} \bar{C}_7: A^3 8B^2C^2, & C_3: A8BC \cdot C_4: A^2 8BCD, & C_3{}^2: A8BCD \cdot C_1: AD, \\ C_3{}^2: A8BCD \cdot C_1': A, & C_3: A8BCD \cdot C_4': A^2 8B, & C_7': A^3 8B^2. \end{array}$$

Let

$$(7)\quad y_1 = \bar{C}_7,\ y_2 = C_3 \cdot C_4,\ y_3 = C_3{}^2 \cdot C_1,\ y_4 = C_3{}^2 \cdot C_1',\ y_5 = C_3 \cdot C_4',\ y_6 = C_7',$$

and project from the line $y_1 = y_2 = y_3 = y_4 = 0$ into the opposite S_3 of the S_5, thus obtaining a surface whose parametric equations are

$$(8)\quad \begin{array}{ll} y_1 = C_7: A^3 8B^2C^2, & y_2 = C_3: A8BCD \cdot C_4: A^2 8BCD, \\ y_3 = C_3{}^2: A8BCD \cdot C_1: 8AD, & y_4 = C_3{}^2: A8BCD \cdot C_1: 8A. \end{array}$$

The surface defined by equations (8) is a rational quartic surface F_4 since the system (8) is of grade 4.

For a general point on $C_3: A8BCD$, $y_1 \neq 0$ and $y_2 = y_3 = y_4 = 0$, hence the image of $C_3: A8BCD$ is the point $(1, 0, 0, 0)$ on F_4. The section of F_4 made by a plane $ky_2 + ly_3 = my_4 = 0$ through the point $(1, 0, 0, 0)$ determines in the plane (x) a composite curve

$$\begin{aligned} C_3: A8BCD(kC_4: A^2 8BCD + lC_3: A8BCD \cdot C_1: AD \\ + mC_3: A8BCD \cdot C_1': A) = 0 \end{aligned}$$

that is

$$(9)\qquad C_3: A8BCD \cdot C_4: A^2 8BCD.$$

Any two curves of form (9) have two residual intersections. Therefore $(1, 0, 0, 0)$ is a double point. Since to each plane section through $(1, 0, 0, 0)$ corresponds a single C_4, hence but one direction through the point D in (x), any plane section through $(1, 0, 0, 0)$ has a cusp at that point.

Let a point P in (x) approach a point Q, other than D, on C_3; then y_3 and y_4 vanish to the second order, y_2 to the first order, and y_1 is finite; hence the image of P approaches $(1, 0, 0, 0)$ along $y_3 = y_4 = 0$, and $y_3 = y_4 = 0$ is the cuspidal tangent. As a point P approaches D on C_3, y_3 vanishes to the

third order, hence $(1, 0, 0, 0)$ is a uniplanar singular point on F_4, the plane of the point being the plane $y_3 = 0$.

The surface (8) is one of the rational quartic surfaces with no double curve, determined by Noether.

4. *The system* $C_9: 8A^3B^2$. In a plane (x) the system of curves $C_9: 8A^3B^2$ with triple points at A_i $(i = 1, 2, \cdots, 8)$ and a double point at B is of dimension 4. Taking $C_3: 8AB$ and $C_3: 8A$ as members of the pencil of cubics on A_i; $C_6: 8A^2B$, a non-composite sextic; and $C_9: 8A^3B^2$ a non-composite curve of order 9, the equation of the system is

$$(10) \qquad a_1C_9: 8A^3B^2 + a_2C_3: 8AB \cdot C_6: 8A^2B + a_3C_3{}^2: 8AB \cdot C_3: 8A + a_4C_3{}^3: 8AB = 0.$$

The cubic C_3 has with C_9 one residual intersection. Call this point C. Let

$$(11) \qquad \begin{array}{ll} y_1 = C_9: 8A^3B^2C, & y_2 = C_3: 8ABC \cdot C_6: 8A^2B, \\ y_3 = C_3{}^2: 8ABC \cdot C_3: 8A, & y_4 = C_3{}^3: 8ABC. \end{array}$$

These are the parametric equations of a rational surface of order 4 in ordinary space. The image of C_3 is the point $(1, 0, 0, 0)$.

For a point near C in (x) the corresponding values of y_1 and y_2 are infinitesimals of the first order and those of y_3 and y_4 are infinitesimals of the second order and the third order respectively. Hence the image of C is the line $y_3 = y_4 = 0$.

For a point near B in (x) the corresponding values of y_1, y_2 and y_3 are infinitesimals of the second order and that of y_4 is an infinitesimal of the third order. Hence the image of B is a conic in the plane $y_4 = 0$.

There is a pencil of curves $C_9: 8A^3B^2C^2$ given by

$$C_3{}^2: 8ABC \cdot C_3: 8A + \gamma C_3{}^3: 8ABC = 0.$$

These go into the sections of F_4 cut by the pencil of planes $y_4 = \gamma y_3$, these sections being composed of rational cubics and the line $y_3 = y_4 = 0$. To the section made by the plane $y_4 = 0$ corresponds $C_3{}^3: 8ABC$ which is of the form $C_9: 8A^3B^3C^3$, hence the section made by this plane is composed of a conic, image of B, and a line image of C, taken twice.

This surface is the second rational quartic surface with no double point determined by Noether.

5. *The system* $C_{12}: 8A^4B^3$. In a plane (x) the system of curves $C_{12}: 8A^4B^3$ is of dimension 4 and the equation of the system may be written

(12) $a_1C_{12}: 8A^4B^3D + a_2C_3: 8ABD \cdot C_9: 8A^3B^2 + a_3C_3^2: 8ABD \cdot C_6: 8A^2B + a_4C_3^3: 8ABD \cdot C_3: 8A + a_5C_3^4: 8ABD = 0$

where D is the residual intersection of $C_3: 8AB$ and $C_{12}: 8A^4B^3$.

Let

(13) $y_1 = C_{12}: 8A^4B^3D, \quad y_2 = C_3: 8ABD \cdot C_9: 8A^3B^2,$
$y_3 = C_3^2: 8ABD \cdot C_6: 8A^2B, \quad y_4 = C_3^3: 8ABD \cdot C_3: 8A, \quad y_5 = C_3^4: 8ABD.$

These are the parametric equations of a rational F_6 in S_4.

Let E be the residual intersection of $C_{12}: 8A^4B^3D = 0$ and $C_9: 8A^3B^2 = 0$. Then through E passes a curve of the pencil $C_3: 8A$ and a non-composite curve of the net $C_6: 8A^2B$. The image of E is a point P on F_6. Project the F_6 from $(0, 0, 0, 0, 1)$ as a center into the opposite S_3 obtaining an F_5 whose parametric equations are

(14) $y_1 = C_{12}: 8A^4B^3DE, \quad y_2 = C_3: 8ABD \cdot C_9: 8A^3B^2E,$
$y_3 = C_3^2: 8ABD \cdot C_6: 8A^2BE, \quad y_4 = C_3^3: 8ABD \cdot C_3: 8AE.$

Each C_{12} of the system goes into a plane section of F_5 which is a C_5 of genus 4, hence a C_5 with two double points. The locus of these double points must be a double conic K on F_5.

The image of point D is the line $y_3 = y_4 = 0$; the image of point E is a line on the surface.

(15) The first polar of a rational surface F_N with respect to a point P not on the surface is a surface of order $N - 1$, containing the double curve on F_N, the curve of contact of the tangent cone to F_N with P as a vertex, and the singular points on F_N. If F_N is a surface whose plane sections are mapped on a plane (x) by a web of curves of order m, then corresponding to the plane sections of F_N through P is a net of curves belonging to the web, whose Jacobian is a curve of order $3(m - 1)$ having a $(3r - 1)$-fold point at an r-fold point of the web. The Jacobian is the image of the curve of contact of the tangent cone with F_N.

For the case at hand the polar surface is an F_4 whose intersection with F_5 goes into a composite $C_{48}: 8A^{16}B^{12}D^4E^4$ consisting of the Jacobian $C_{33}: 8A^{11}B^8D^2E^2$; the curve $C_3: 8ABD$, image of $(0, 0, 0, 1)$; and the image of the double conic K which is a $C_{12}: 8A^4B^3DE^2$. A general $C_{12}: 8A^4B^3DE$ goes into a general plane section of F_5 but the $C_{12}: 8A^4B^3DE^2$ goes into the double conic counted twice and a residual line of F_5 in the plane of the conic. This line is the image of point E.

5(a). If point B is chosen on a certain locus there is a $C_9: 8A^3B^3$ other than

$C_3^3: 8AB$.* Hence we can choose as linearly independent members of the system

$$(16)\quad \begin{array}{lll} C_9: 8A^3B^3 \cdot C_3: 8ABC, & C_9: 8A^3B^3 \cdot C_3: 8AC, & \\ C_3^3: 8ABC \cdot C_3: 8AC, & C_3^2: 8ABC \cdot C_6: 8A^2B, & C_3^4: 8ABC, \end{array}$$

where C is the ninth point common to $C_3: 8ABC$ and $C_3: 8A$.

Take any point D in the plane (x). There is a member of the pencil $C_9: 8A^3B^3 + \gamma C_3^3: 8ABC = 0$ through D. Call this a new $C_9: 8A^3B^3$. There is also a $C_3: 8A$ and a $C_6: 8A^2B$ through D.

Let

$$(17)\quad \begin{array}{ll} y_1 = C_9: 8A^3B^3D \cdot C_3: 8ABC, & y_2 = C_9: 8A^3B^3D \cdot C_3: 8ACD, \\ y_3 = C_3^2: 8ABC \cdot C_6: 8A^2BD, & y_4 = C_3^3: 8ABC \cdot C_3: 8ACD. \end{array}$$

These are the parametric equations of a rational surface F_5 in S_3.†

The image of $C_9: 8A^3B^3D$ is the line $y_1 = y_2 = 0$. The image of $C_3: 8ACD$ is the line $y_2 = y_4 = 0$. The image of the point C is the line $y_3 = y_4 = 0$, the image of point D a line in the plane $y_2 = 0$, and the image of point B a rational cubic in the plane $y_1 = 0$, with a double point at $(0, 1, 0, 0)$.

Since C_9 has two residual intersections with a general C_{12}, each plane section of F_5 has a double point on the line $y_1 = y_2 = 0$, the image of C_9. Hence $y_1 = y_2 = 0$ is a double line.

A section of F_5 made by a plane of the pencil $y_2 = \gamma y_4$ goes into a composite C_{12} of the form $C_3: 8ACD \cdot C_9: 8A^3B^3$, which meets a general C_{12} in 3 residual points not on $C_3: 8ACD$. Hence the line of intersection of any plane of the pencil and a plane not of the pencil meets F_5 in 3 points other than on the line $y_2 = y_4 = 0$, which means that $y_2 = y_4 = 0$ is a double line.

We observe that the surface F_5 as defined by (17) has a composite double conic consisting of the double lines $y_1 = y_2 = 0$ and $y_2 = y_4 = 0$. As in the general case (14), the image line of the point D is the residual intersection of the plane of the double conic with F_5.

6. *The system* $C_{15}: 8A^5B^4$. In a plane (x) the system of curves $C_{15}: 8A^5B^4$ is of dimension 5 and by the method employed in the previous cases we can choose as linearly independent members of the system

$$(18)\quad \begin{array}{lll} C_{15}: 8A^5B^4, & C_{12}: 8A^4B^3 \cdot C_3: 8AB, & C_9: 8A^3B^2 \cdot C_3^2: 8AB, \\ C_6: 8A^2B \cdot C_3^3: 8AB, & C_3^4: 8AB \cdot C_3: 8A, & C_3^5: 8AB. \end{array}$$

---

* Halphen, *Bulletin de la So. Math.*, 162 (1882).

† D. Montesano, *Rendiconti della Reale Accodemia di Napoli*, Ser. 3, Vol. 13 (1907), pp. 66-68.

Let C be the residual intersection of $C_3: 8AB$ and $C_{15}: 8A^5B^4$. The curves $C_{15}: 8A^5B^4$ and $C_{12}: 8A^4B^3$ have 8 residual points of intersection. Call two of these points D and E. Choose a new C_9 and a new C_6 such that they will pass through D and E.

Let

$$
(19)\quad \begin{aligned}
&y_1 = C_{15}: 8A^5B^4CDE, && y_2 = C_{12}: 8A^4B^3DE \cdot C_3: 8ABC,\\
&y_3 = C_9: 8A^3B^2DE \cdot C_3{}^2: 8ABC, && y_4 = C_6: 8A^2BDE \cdot C_3{}^3: 8ABC,\\
&y_5 = C_3{}^4: 8ABC \cdot C_3 8AD, && y_6 = C_3{}^5: 8ABC.
\end{aligned}
$$

These are the parametric equations of a rational F_8 in S_5. The images of points D and E are points on the line $y_1 = y_2 = y_3 = y_4 = 0$.

From the line $y_1 = y_2 = y_3 = y_4 = 0$ as a center project F_8 into the S_3, $y_5 = y_6 = 0$, giving a surface whose parametric equations are

$$
(20)\quad \begin{aligned}
&y_1 = C_{15}: 8A^5B^4CDE, && y_2 = C_{12}: 8A^4B^3DE \cdot C_3: 8ABC,\\
&y_2 = C_9: 8A^3B^2DE \cdot C_3{}^2: 8ABC, && y_4 = C_6: 8A^2BDE \cdot C_3{}^3: 8ABC.
\end{aligned}
$$

The surface is an F_6. The points D and E go into lines on this surface. The image of C is the line $y_3 = y_4 = 0$. The image of $C_3: 8ABC$ is the point $(1, 0, 0, 0)$.

Since the genus of a member of the system is 5, each plane section of F_6 is a C_6 of genus 5. Hence there is a double C_5 on the surface.

By (15) we see that the section of F_6 by the first polar of a point not on F_6 goes into a composite $C_{75}: 8A^{25}B^{20}C^5D^5E^5$ which consists of the Jacobian $C_{42}: 8A^{14}B^{11}C^2D^2E^2$; $C_3: 8ABC$, the image of the singular point $(0, 0, 0, 1)$; and the image of the double C_5 which is a $C_{30}: 8A^{10}B^8C^2D^3E^3$. A general $C_{30}: 8A^{10}B^8C^2D^2E^2$ goes into a general quadric section of F_6 but the $C_{30}: 8A^{10}B^8C^2D^3E^3$ goes into the quadric section containing the double C_5 with the image lines of points D and E as the residual intersection of the quadric with F_5.

If two surfaces of order n_1 and n_2 respectively contain a C_m of order m and rank r, genus p to multiplicity i_1 and i_2 respectively, the residual C_v meets C_m in t points and has genus π where *

$$
(21)\quad \begin{aligned}
&t = m(i_2n_1 + i_1n_2 - 2i_1i_2) - i_1i_2r,\\
&\pi = [v(n_1 + n_2 - 4) - (i_1 + i_2 - 1)t]/2 + 1,\\
&r = 2m + 2p - 2.
\end{aligned}
$$

---

* Noether, *Annali di Mathematica*, Ser. 2, Vol. 5 (1871), pp. 163-177.

For the case in question

$$\begin{array}{lll} n_1 = 2 & i_1 = 1 & m = 5 \\ n_2 = 6 & i_2 = 2 & \end{array}$$

and, since C_v is a composite conic,

$$v = 2 \text{ and } \pi = -1.$$

Substituting the above values in equations (21) we find that $p = 2$; that is, the double C_5 on F_6 is a curve of genus 2.

A C_5 genus 2 is the partial intersection of a quadric and a cubic surface, the residual being a ruling on the quadric.

Let F_3 be a cubic surface containing the C_5 and one line of the degenerate C_2. Then the number of intersections of the line with C_5 is, by (21), $t = 3$. Hence the line images of points D and E are trisecants of the quintic C_5.

6(a). If B is chosen on a certain locus there is a $C_{12}: 8A^4B^4$ other than $C_3^4: 8AB$, and we can take as parametric equations of the surface

$$(22) \quad \begin{array}{ll} y_1 = C_{12}: 8A^4B^4DE \cdot C_3: 8AC, & y_2 = C_{12}: 8A^4B^4DE \cdot C_3: 8ABC, \\ y_3 = C_9: 8A^3B^2DE \cdot C_3^2: 8ABC, & y_4 = C_6: 8A^2BDE \cdot C_3^3: 8ABC. \end{array}$$

The surface is again an F_6 with plane sections of genus 5 and a double curve C_5 of order 5.

The image of C_{12} is the line $y_1 = y_2 = 0$. Since $C_{12}: 8A^4B^4DE$ meets a general $C_{15}: 8A^5B^4CDE$ in 2 residual points, $y_1 = y_2 = 0$ is a double line; hence the double C_5 is composite with $y_1 = y_2 = 0$ as a component.

For a point near B in (x), y_1, y_3 and y_4 are infinitesimals of order four while y_2 is an infinitesimal of order five, hence the image of B is a rational quartic in the plane $y_2 = 0$. The images of D and E are lines meeting the line $y_1 = y_2 = 0$, since D and E are on C_{12} in plane (x).

6(b). Again if B is chosen on a certain locus so that there is a $C_9: 8A^3B^3$ other than $C_3^3: 8AB$ we can take the following as parametric equations of a surface

$$(23)^* \quad \begin{array}{l} y_1 = C_9: 8A^3B^3DEF \cdot C_6: 8A^2BCDEF, \\ y_2 = C_9: 8A^3B^3DEF \cdot C_3: 8ABC \cdot C_3: 8A, \\ y_3 = C_9: 8A^3B^3DEF \cdot C_3^2: 8ABC, \\ y_4 = C_6: 8A^2BDEF \cdot C_3^3: 8ABC. \end{array}$$

The surface is an F_5 genus 5. Since each plane section is a quintic curve of genus 5 there is a double line on the surface.

---

* D. Montesano, *Rendiconti della Reale Accademia di Napoli*, Ser. 3, Vol. 13 (1907), pp. 66-68.

The image of C_9 is the point $(0,0,0,1)$. The image of $C_3: 8AB$ is the point $(1,0,0,0)$. The image of C is the line $y_3 = y_4 = 0$. The images of points D, E and F which are on C_9 are three lines in the plane $y_1 = 0$, passing thru $(0,0,0,1)$, the image of C_9.

The image of C_6 is the line $y_1 = y_4 = 0$. Since $C_6: 8A^2BCDEF$ has, with a general $C_{15}: 8A^5B^4CDEF$, two residual intersections, every plane section of F_5 has a double point on $y_1 = y_4 = 0$, which is, then, the double line on the surface.

A plane section thru the point $(0,0,0,1)$ goes into a composite curve in plane (x) of the form

$$C_9: 8A^3B^3DEF \cdot C_6: 8A^2BC.$$

Two curves of this type meet in two residual points not on C_9, hence the line of intersection of two planes through $(0,0,0,1)$ meets F_5 in two points other than $(0,0,0,1)$, hence $(0,0,0,1)$ is a triple point on F_5.

By the method of (15) we can show that the images of the points D, E and F in (x) form the residual intersection of the plane through the double line and the point $(0,0,0,1)$.

7. *The system* $C_{18}: 8A^6B^5$. This system is of dimension 6 and by a process of reasoning similar to that in the previous cases we obtain the surface in S_3 given parametrically by

$$(24)\quad \begin{aligned} &y_1 = C_{18}: 8A^6B^5CD^2, && y_2 = C_{15}: 8A^5B^4D^2 \cdot C_3: 8ABC, \\ &y_3 = C_{12}: 8A^4B^3D^2 \cdot C_3^2: 8ABC, && y_4 = C_9: 8A^3B^2D^2 \cdot C_3^3: 8ABC, \end{aligned}$$

The surface is an F_6 and each plane section is a C_6 genus 5; hence the surface contains a double C_5 of order 5. The image of the curve C_3 is the point $(1,0,0,0)$. The image of the point C is the line $y_3 = y_4 = 0$. The image of the point D is a conic on F_6.

Through the double C_5 on F_6 passes one quadric surface whose residual intersection with F_6 is a conic C_2.

Again referring to (15) we can show that the double curve C_5 goes into a $C_{36}: 8A^{12}B^{10}C^2D^5$. Since a general $C_{36}: 8A^{12}B^{10}C^2D^4$ goes into a general quadric section of F_6, the $C_{36}: 8A^{12}B^{10}C^2D^5$ goes into the section made by the quadric on the double C_5, which, therefore, has the image conic of point D as the residual intersection with F_6.

7(a). If B is chosen so that there is a $C_9: 8A^3B^3$ other than $C_3^3: 8AB$ we obtain a surface whose parametric equations are

$$(25)^* \quad \begin{aligned} y_1 &= C_9: 8A^3B^3DE \cdot C_9: 8A^3B^2CD, \\ y_2 &= C_9: 8A^3B^3DE \cdot C_6: 8A^2BD \cdot C_3: 8ABC, \\ y_3 &= C_9: 8A^3B^3DE \cdot C_3: 8AD \cdot C_3{}^2: 8ABC, \\ y_4 &= C_{12}: 8A^4B^3D^2E \cdot C_3{}^2: 8ABC. \end{aligned}$$

The surface is an F_5 with plane sections of genus 5; hence there is a double line on the surface. The image of $C_9: 8A^3B^3$ is the point $(0, 0, 0, 1)$. The image of $C_3: 8ABC$ is the point $(1, 0, 0, 0)$. The images of D and E are a conic and a line, both of which must pass through the point $(0, 0, 0, 1)$, since D and E are on $C_9: 8A^3B^3$. The image of C is the line $y_3 = y_4 = 0$.

The plane determined by the double line and the point $(0, 0, 0, 1)$ has a C_3 as residual intersection with F_5.

From (15) we find that the double line goes into a $C_{18}: 8A^6B^5CD^3E^2$. Since a general C_{18} goes into a general plane section of F_5, the $C_{18}: 8A^6B^5CD^3E^2$ goes into the section made by the plane through the double line and the point $(0, 0, 0, 1)$ and contains the images of points D and E as the residual intersection with F_5. The residual C_3 in which the plane determined by the double line and the point $(0, 0, 0, 1)$ meets the F_5 is therefore composite, and consists of a conic and a line, images of D and E.

A section of F_5 by a plane $a_1y_1 + a_2y_2 + a_3y_3 = 0$ goes into a composite curve of the form

$$C_9: 8A^3B^3DE(a_1C_9: 8A^3B^2CD + a_2C_6: 8A^2BD \cdot C_3: 8ABC \\ + a_3C_3: 8AD \cdot C_3: 8ABC) = 0;$$

that is

$$C_9: 8A^3B^3DE \cdot C_9: 8A^3B^2CD.$$

Two of these $C_9: 8A^3B^2CD$ meet in three residual points, hence $(0, 0, 0, 1)$ is a double point on F_5.

7(b). If the point B is so chosen that there is a $C_{15}: 8A^5B^5$ other than $C_3{}^5: 8AB$, the parametric equations of the surface may be taken as

$$(26) \quad \begin{aligned} y_1 &= C_{15}: 8A^5B^5DEF \cdot C_3: 8AC, & y_2 &= C_{15}: 8A^5B^5DEF \cdot C_3: 8ABC, \\ y_3 &= C_{12}: 8A^4B^3DEF \cdot C_3{}^2: 8ABC, & y_4 &= C_9: 8A^3B^2DEF \cdot C_3{}^3: 8ABC. \end{aligned}$$

This surface differs from (24) only in that the line $y_1 = y_2 = 0$ is a double line on the surface and a component of the double C_9.

7(c). If the point B is so chosen that there is a $C_{12}: 8A^4B^4$ other than $C_3{}^4: 8AB$ then we obtain a surface whose parametric equations are

* D. Montesano, *Rendiconti della Reale Accademia di Napoli*, Ser. 3, Vol. 13 (1907), pp. 66-68.

$$(27)\quad \begin{aligned} y_1 &= C_{12}: 8A^4B^4DEFG \cdot C_6 8A^2BC, \\ y_2 &= C_{12}: 8A^4B^4DEFG \cdot C_3: 8ABC \cdot C_3: 8A, \\ y_3 &= C_{12}: 8A^4B^4DEFG \cdot C_3^2: 8ABC, \\ y_4 &= C_9: 8A^3B^2DEFG \cdot C_3^3: 8ABC. \end{aligned}$$

The surface is an F_6 with plane sections of genus 6 and contains a double C_4. The image of C_{12} is the point $(0, 0, 0, 1)$. A section made by a plane through this point goes into a curve in plane (x) of the form

$$C_{12}: 8A^4B^4DEFG \cdot C_6: 8A^2BC.$$

Two of these $C_6: 8A^2BC$ meet in two residual points, hence the point $(0, 0, 0, 1)$ is 4-fold on the surface.

The images of D, E, F, and G are lines in the plane $y_1 = 0$ passing through $(0, 0, 0, 1)$. The image of C_6 is the line $y_1 = y_4 = 0$. In plane (x), $C_6: 8A^2BCDEFG$ meets a general $C_{18}: 8A^6B^5CDEFG$ in two residual points, hence $y_1 = y_4 = 0$ is a double line and a component of the double C_4.

From (15) we find that the double C_4 goes into a $C_{36}: 8A^{12}B^{10}C^2D^3E^3F^3G^3$. This C_{36} goes into the quadric section on the double C_4 and the point $(0, 0, 0, 1)$ which has as residual intersection with F_6 the composite C_4 consisting of the four lines, images of the points D, E, F and G.

8. *The system* $C_{21}: 8A^7B^6$. This system is of dimension 7 and by the methods previously employed we obtain the surface whose parametric equations are

$$(28)\quad \begin{aligned} y_1 &= C_{21}: 8A^7B^6CD^2E, & y_2 &= C_{18}: 8A^6B^5D^2E \cdot C_3: 8ABC, \\ y_3 &= C_{15}: 8A^4B^3D^2E \cdot C_3^2: 8ABC, & y_4 &= C_{12}: 8A^3B^2D^2E \cdot C_3^3: 8ABC. \end{aligned}$$

The surface is an F_7 with plane sections of genus 6, hence there is a double curve C_9 of order 9 on the surface.

The image of C is the line $y_3 = y_4 = 0$, the image of D a conic and the image of E a line. By (15) we find that the double C_9 goes into a curve in the plane (x) of the form $C_{63}: 8A^{21}B^{18}C^3D^7E^3$. This C_{63} goes into the section of F_7 made by the cubic surface containing the double C_9 and has as residual intersection a composite cubic curve consisting of the conic image of D and the line image of E.

8(a). If B is so chosen that there is a $C_{15}: 8A^5B^5$ other than $C_3^5: 8AB$ a surface is determined whose parametric equations are

$$(29)\quad \begin{aligned} y_1 &= C_{15}: 8A^5B^55D \cdot C_6: 8A^2BC, & y_2 &= C_{15}: 8A^5B^55D \cdot C_3: 8ABC \cdot C_3 8A, \\ y_3 &= C_{15}: 8A^5B^55D \cdot C_3^2: 8ABC, & y_4 &= C_{12}: 8A^4B^35D \cdot C_3^3: 8ABC. \end{aligned}$$

The surface is an F_7 with plane sections of genus 7 hence there is a double curve C_8 of order 8 on the surface. The image of C is the line $y_3 = y_4 = 0$. The image of $C_3: 8ABC$ is the point $(0, 0, 0, 1)$. Since a line through $(0, 0, 0, 1)$ meets the surface in only two residual points this point is of order 5 on the surface. The image of B is a rational sextic in the plane $y_3 = 0$ with a point of order 5 at $(0, 0, 0, 1)$.

The images of the D_i $(i = 1, 2, \cdots, 5)$ are five lines on the surface passing through the point $(0, 0, 0, 1)$ and form the residual intersection with F_7 of the cubic surface through the double C_8.

8(b). If the point B is so chosen that there is a $C_{12} 8A^4B^4$ other than $C_3{}^4: 8AB$ a surface is determined whose parametric equations are

$$(30) \quad \begin{aligned} y_1 &= C_{12}: 8A^4B^4DEF \cdot C_9: 8A^3B^2CD^2EF, \\ y_2 &= C_{12}: 8A^4B^4DEF \cdot C_6: 8A^2BD \cdot C_3: 8ABC, \\ y_3 &= C_{12}: 8A^4B^4DEF \cdot C_3{}^2: 8ABC \cdot C_3: 8AD, \\ y_4 &= C_9: 8A^3B^2CD^2EF \cdot C_3{}^4: 8ABC. \end{aligned}$$

The surface is an F_6 with plane sections of genus 6, hence there is a double C_4 on the surface.

The image of C_{12} is the point $(0, 0, 0, 1)$, the image of $C_3: 8ABC$ the point $(1, 0, 0, 0)$, and the image of C_9 is the line $y_1 = y_4 = 0$. The image of D is a conic in the plane $y_1 = 0$ passing through the point $(0, 0, 0, 1)$. The images of E and F are lines in the plane $y_1 = 0$. The line $y_1 = y_4 = 0$ is a double line. The point $(0, 0, 0, 1)$ is a triple point on the surface.

By the methods employed in the previous cases we find that the residual intersection with F_6 of the quadric through the double C_4 is a composite C_4 consisting of the conic image of the point D and the line images of the points E and F.

9. *Conclusion.* It is now clear that the processes developed in this paper can be carried on indefinitely for any linear system of the type $C_{3n}: 8A^nB^{n-1}$ containing a pencil $C_{3i}: 8A^iB^i$.

A Problem of Ambience.

By William Kelso Morrill.

In the following paper, we shall consider a triangle of directed lines, the vertices of which are moving with the same constant speed parallel to their respective opposite sides but in opposite directions. The invariants of the triangle are studied, and the motions of the vertices are investigated by the aid of the Weierstrass elliptic function theory as well as the q-series of Jacobi.

1. *The Invariants of the Triangle.* Let a, b, c be the lengths of the sides of the triangle, and α, β, γ their respective directions. If θ, ϕ, ψ are the angles which the sides of the triangle make with the base line, then $-\alpha = e^{i\theta}$; $\beta = e^{i\phi}$; $-\gamma = e^{i\psi}$. We shall use A, B, C in two senses: first as the affices of the vertices, second as the interior angles of the triangle. The motion of the vertices is given by the following diffrential equations

$$\dot{A} = -\alpha v, \quad \dot{B} = -\beta v, \quad \dot{C} = -\gamma v,$$

where the dot indicates differentiation with respect to the time and v is the speed. We can now write $a\alpha = C - B$. Then

$$D_t a\alpha \equiv \dot{a}\alpha + \dot{\alpha}a = \dot{C} - \dot{B} = (\beta - \gamma)v$$

and
$$\dot{a} + (\dot{\alpha}/\alpha)a = (\beta/\alpha - \gamma/\alpha)\, v.$$

But
$$\beta/\alpha = e^{i(\pi - C)} = -\cos C + i \sin C$$

and
$$\gamma/\alpha = e^{i(\pi + B)} = -\cos B - i \sin B.$$

$$\therefore \quad \dot{a} + a\dot{\alpha}/\alpha = v[\cos B - \cos C + i(\sin B + \sin C)].$$

Equating reals and imaginaries,

$$\dot{a} = (\cos B - \cos C)v,$$
$$a\dot{\alpha}/\alpha = i(\sin B + \sin C)v.$$

The variations of the sides of the triangle are given, therefore, by

$$1.1 \quad \begin{cases} \dot{a} = (\cos B - \cos C)v, \\ \dot{b} = (\cos C - \cos A)v, \\ \dot{c} = (\cos A - \cos B)v. \end{cases}$$

Now $v = d\Lambda/dt$, where Λ is the distance each body moves in the time t. Choosing $v = 1$, we have $d\Lambda = dt$ and $\Lambda = t$. Adding equations 1.1, we have $\dot{a} + \dot{b} + \dot{c} = 0$.

1.2 $$\therefore \qquad a + b + c = s_1$$

where s_1 is a constant; and our first result is that the perimeter of the triangle remains constant. Finding the perimeter constant suggests a study of the area.

$$\text{Area} = [s(s-a)(s-b)(s-c)]^{1/2},$$

where $s = (a+b+c)/2$. Thus, letting

$$X \equiv 16(\text{Area})^2 = 2(a^2b^2 + b^2c^2 + a^2c^2) - a^4 - b^4 - c^4$$
$$\dot{X} = 4\{(b^2 + c^2 - a^2)\dot{a}a + (c^2 + a^2 - b^2)\dot{b}b + (a^2 + b^2 - c^2)\dot{c}c\}$$
$$= 8abc(\dot{a}\cos A + \dot{b}\cos B + \dot{c}\cos C).$$

Substituting the values of $\dot{a}$, $\dot{b}$, $\dot{c}$ from 1.1; we have: $\dot{X} = 0$ and

1.3 $X = S_3$, where S_3 is a constant; that is, the area also is constant.

We thus find *the perimeter and area of the triangle are invariant under the motion.*

2. *Introducing the Elliptic Functions.* Let $x = s - a$, $y = s - b$, $z = s - c$. Then from 1.2 and 1.3 respectively, we obtain

$$x + y + z = k_1, \qquad \text{and}$$
$$xyz = k_3.$$
$$\therefore \qquad dx + dy + dz = 0, \qquad \text{and}$$
$$yz\,dx + xz\,dy + xy\,dz = 0.$$

From these two equations, we obtain

2.1 $$dx/x(y-z) = dy/y(z-x) = dz/z(x-y) = d\mu,$$
$$\therefore dz/d\mu = z(x-y) = z\{(x+y)^2 - 4xy\}^{1/2} = z[(k_1 - z)^2 - 4k_3/z]^{1/2}$$
$$(dz/d\mu)^2 = z^2[(k_1 - z)^2 - 4k_3/z].$$

Put $z = -1/v$; then $dz/d\mu = v^{-2}dv/d\mu$, and

$$(dv/d\mu)^2 = (k_1v + 1)^2 + 4k_3v^3.$$

Now putting $v = w - k_1^2/12k_3$, we get

2.2 $$(dw/d\mu)^2 = 4k_3w^3 + (2k_1 - k_1^4/12k_3)w + (1 - 2k_1^3/12k_3 + 2k_1^6/3\cdot 12^2k_3^2).$$

Finally putting $k_3^{1/2}d\mu = du$, and

2.3 $$g_2 = (-2k_1/k_3)(1 - k_1^3/24k_3) \quad \text{and}$$

2.4 $$g_3 = (-1/k_3)(k_1^6/216k_3^2 - k_1^3/6k_3 + 1) \quad \text{and}$$

substituting in 2.2 we obtain

2.5 $$(dw/du)^2 = 4w^3 - g_2 w - g_3.$$

Equation 2.5 is the elliptic relation $(p'u)^2 = 4p^3u - g_2pu - g_3$, and hence our problem is an elliptic function problem.

We may then write $w = p(u - \gamma)$, where γ is a constant. Since $z = -1/v$, and $v = w - k_1^2/12k_3$, $z = 1/[k_1^2/12k_3 - p(u-\gamma)]$. Setting

2.6 $$k_1^2/12k_3 = p\tau$$

and noting, from 2.1, that x and y have expressions similar to z, we have:

2.7 $$\begin{cases} x = 1/[p\tau - p(u-\alpha)], \\ y = 1/[p\tau - p(u-\beta)], \\ z = 1/[p\tau - p(u-\gamma)]. \end{cases}$$

Note that $z^{-1}dz/du = (x-y)/k_3^{1/2} = p'(u-\gamma)/[p\tau - p(u-\gamma)]^2$. This equals zero when $x = y$, which is at the half periods of the parallelogram of periods, since we know the function p' is zero there. Conversely when $u - \gamma$ is a half period, that is when

2.8 $$u - \gamma \equiv m_1\omega_1 + m_2\omega_2,$$ where

$m_1, m_2 = 0, 1, 2$ but $m_1 = m_2 \neq 0$ or 2, $p'(u-\gamma) = 0$, and $x = y$ or from 2.7, $p(u-\alpha) = p(u-\beta)$.

Consider then $u - \alpha \equiv \beta - u$, and, therefore, $2u \equiv \alpha + \beta$. Since from 2.8 we have $2u \equiv 2\gamma$, it follows that

2.9 $$\alpha + \beta \equiv 2\gamma, \quad \beta + \gamma \equiv 2\alpha, \quad \gamma + \alpha \equiv 2\beta.$$
$$\therefore \quad \alpha - \gamma \equiv 2\gamma - 2\alpha \quad \text{or} \quad 3\alpha \equiv 3\beta \equiv 3\gamma.$$

This result tells us that α, β, γ in 2.7 are constants which differ from each other by thirds of a period. By choosing $\gamma = 0$, it follows that $\alpha = \alpha$, and $\beta = 2\alpha$, and we can rewrite 2.7 in the following way:

2.7′ $$\begin{cases} x = 1/(p\tau - pu), \\ y = 1/[p\tau - p(u+\alpha)], \\ z = 1/[p\tau - p(u-\alpha)]. \end{cases}$$

x has poles at $u = \pm\tau$; y has poles at $u = \pm\tau - \alpha$; and z has poles at $u = \pm\tau + \alpha$. Since $x + y + z$ is a constant, the sum of the poles of x, y, and z respectively must be a period. Hence τ must be a third of a period.

The type of network can now be determined quite easily: g_2 and g_3 are both real and the discriminant

$$\Delta \equiv g_2^3 - 27g_3^2 = (k_1^3 - 27k_3)/k_3^3 > 0.$$

This follows from the theorem: *If n numbers $x_1, \cdots x_n$ are positive, the arithmetical mean must be equal to or greater than the geometrical mean.** Since g_2 and g_3 are real and $\Delta > 0$, our net work is rectangular.

We are interested in how the triangle behaves as the elliptic parameter u moves in a rectangular cell. But there are limitations on how u shall move in the cell.

There are eight thirds of a period in a cell. Of these, only four give distinct values to pu, due to the evenness of the p function. At the vertices of the cell pu is infinite. Along the boundaries it is real. As we move along the rectangle of half periods, pu decreases from $+\infty$ to $-\infty$ and is real. u must vary along such a path as will keep k_1 and k_3 real and positive. To find k_1 and k_3 in terms of elliptic functions, we proceed as follows

$$x = 1/(p\tau - pu) = 1/[p\tau - 1/u^2 - c_2u^2 \cdots]$$

and this equals zero for $u = 0$.

Expanding $p(u+\tau)$ in a Taylor's series we obtain for y the following:

$$\begin{aligned} y &= 1/[p\tau - p(u+\tau)] = 1/[p\tau - (p\tau + up'\tau + u^2p''\tau/2! + \cdots)] \\ &= 1/[-up'\tau(1 + up''\tau/2!\,p'\tau + u^2p'''\tau/3!\,p'\tau + \cdots)] \\ &= (-1/up'\tau)[1 - up''\tau/2!\,p'\tau - u^2p'''\tau/3!\,p'\tau + \cdots]. \end{aligned}$$

In a similar way we obtain

$$z = (1/up'\tau)[1 + up''\tau/2!\,p'\tau - u^2p'''\tau/3!\,p'\tau + \cdots];$$

whence it follows

$$x + y + z = p''\tau/p'^2\tau + \text{a function of } u.$$

when $u = 0$, this function vanishes. Hence since $x + y + z = k_1$ is a constant, we have

$$k_1 = p''\tau/p'^2\tau.$$

From 2.6 we have $k_3 = k_1^2/12p\tau = p''^2\tau/12p'^4\tau \cdot p\tau$. Since τ is a third of a period, we have $12p\tau p'^2\tau = p''^2\tau$; for from the identity $2pu + p(2u) = p''^2u/4p'^2u$, we obtain on letting $u = \tau$,

$$2p\tau + p(2\tau) = p''^2\tau/4p'^2\tau.$$

But $p\tau = p(2\tau)$, and hence $12p\tau \cdot p'^2\tau = p''^2\tau$. Putting this back in the expression for k_3 above we obtain

* Todhunter's *Algebra*, Page 422.

$$k_3 = 1/p'^2\tau.$$

To keep k_1 and k_3 real and positive, $p'\tau$ must be real, and $p''\tau$ must be real and positive. Both of these conditions hold when τ lies on the real axis. Furthermore the path along which u moves on the cell must keep 2. 7′ positive and real. There is only one choice: u must move along the path which joins the mid-points of the vertical sides of the cell. Thus in our problem $u = \omega_2 + v$, where v is real and varies from 0 to $2\omega_1$.

3. *The Isosceles Cases.* Knowing the path along which u must move, we will next determine when the triangle becomes isosceles. Let us consider first $x = y$. Then $p(u + \tau) = pu$, hence $u + \tau = \pm u + 2m_1\omega_1 + 2m_2\omega_2$. The case of interest here is the one leading to $2u + \tau = 2m_1\omega_1 + 2m_2\omega_2$; but $u = \omega_2 + v$,

$$\therefore \quad 2\omega_2 + 2v + \tau = 2m_1\omega_1 + 2m_2\omega_2;$$

$$\begin{aligned} &\text{thus} & 2v + \tau &= 0, & &\text{whence} & v &= -\tau/2 = 5\tau/2, \\ &\text{or} & &= 3\tau, & &\text{whence} & v &= \tau, \\ &\text{or} & &= 6\tau, & &\text{whence} & v &= 5\tau/2, \\ &\text{or} & &= 9\tau, & &\text{whence} & v &= 4\tau = \tau. \end{aligned}$$

Hence the sides a and b of the triangle become equal for two values of u; viz.,

$$u = 5\tau/2 + \omega_2 \quad \text{and} \quad u = \tau + \omega_2.$$

Next we will consider $x = z$. Then $p(u - \tau) = pu$ and going through a similar argument we find the sides a and c of our triangle are equal when

$$u = \omega_2 + \tau/2, \quad \text{and} \quad u = \omega_2 + 2\tau.$$

Finally we have $y = z$ when $p(u + \tau) = p(u - \tau)$. In this case we find that

$$u = \omega_2 \quad \text{and} \quad u = \omega_2 + 3\tau/2,$$

which tells us that our triangle was initially isosceles as $b = c$.

We can sum up the results of this section thus: Starting isosceles, the triangle becomes isosceles at every sixth of a period as u moves along the path $u = \omega_2 + v$, starting with $v = 0$.

4. *The Positional Equations.* We shall determine Λ as a function of the elliptic parameter u. From 1. 1 we have

$$da/d\Lambda = \cos B - \cos C = (-2k_1/abc)(s-a)[(s-c)-(s-b)]$$

$$\therefore \quad da/d\Lambda = (-2k_1k_3/abc)[p'up'\tau/(pu - p\tau)^2].$$

Now $s-a=1/(p\tau-pu)$. Therefore $da/du=-p'u/(p\tau-pu)^2$, and $d\Lambda/du=abc/2k_1k_3^{1/2}$; whence

$$d\Lambda/du=-k_3^{1/2}/2k_1+k_3^{1/2}k_1^2/8k_3-(k_3^{1/2}/2)[pu+p(u+\tau)+p(u-\tau)].$$

Calling $K=-k_3^{1/2}/2k_1+k_3^{1/2}k_1^2/8k_3$, we have

$$d\Lambda=Kdu-(k_3^{1/2}/2)[pu+p(u+\tau)+p(u-\tau)]du.$$

4.1 $$\therefore\quad \Lambda=Ku+(k_3^{1/2}/2)[\zeta u+\zeta(u+\tau)+\zeta(u-\tau)]+c.$$

To determine the constant of integration, let $v=0$, then $u=\omega_2$ and $c=K\omega_2+3k_3^{1/2}\eta_2/2$.

Thus for a particular position of u along its path, we can determine the distance the affices have moved.

As the affices move, the rates of change of the angles θ, ϕ, and ψ are given by a set of equations (see Section 1) which we will call positional equations:

4.2 $$\begin{cases} ad\theta/d\Lambda=\sin B+\sin C, \\ bd\phi/d\Lambda=\sin C+\sin A, \\ cd\psi/d\Lambda=\sin A+\sin B. \end{cases}$$

Expressing these in terms of elliptic functions, we have

$$\begin{aligned} d\theta/d\Lambda &=(\sin B+\sin C)/a \\ &=[2(k_1k_3)^{1/2}/abc]\,[(b+c)/a]. \end{aligned}$$

We have already found

$$d\Lambda/du=abc/2k_1k_3^{1/2}$$

Hence
$$\begin{aligned} d\theta/du &=(1/k_1^{1/2})[(b+c)/a] \\ &=(1/k_1^{1/2})\{1+2/[k_1(p\tau-pu)-1]\}. \end{aligned}$$

If we make the substitution

4.3 $$pv_0=p\tau-1/k_1,$$

where pv_0 is a constant, we get $d\theta=(1/k_1^{1/2})[1+2/k_1(pv_0-pu)]du$.*

Multiplying both sides by $p'v_0$ and integrating, we obtain

$$\theta p'v_0=(1/k_1^{1/2})[up'v_0+\frac{2}{k_1}\,(\log\frac{\sigma(u+v_0)}{\sigma(u-v_0)}-2u\zeta v_0)]+C_1.$$

But, by putting for pv_0 its value given in 4.3, in the identity

$$p'v_0=4p^3v_0-g_2pv_0-g_3$$

we find that $p'v_0=2i/k_1^{3/2}$.

* Halphen, *Traité des Fonctions Elliptiques*, Vol. 1, p. 185.

The three angles of the triangle are then given by the following equations:

$$
4.4 \quad \begin{cases} i\theta = iu/k_1^{1/2} + \log \dfrac{\sigma(u+v_0)}{\sigma(u-v_0)} - 2u\zeta v_0 + C_1, \\[2ex] i\phi = i(u+\tau)/k_1^{1/2} + \log \dfrac{\sigma(u+\tau+v_0)}{\sigma(u+\tau-v_0)} - 2(u+\tau)\zeta v_0 + C_2, \\[2ex] i\psi = i(u-\tau)/k_1^{1/2} + \log \dfrac{\sigma(u-\tau+v_0)}{\sigma(u-\tau-v_0)} - 2(u-\tau)\zeta v_0 + C_3. \end{cases}
$$

The equations 4.4 are important since they tell us the position of the triangle for a particular value of the parameter u.

5. *Introducing the q-series.* The representation of the elliptic functions by the q-series was invented by Jacobi * and is most important for practical problems. His invention made it possible to express doubly periodic functions in an infinite series, the terms of which are singly periodic functions. The problem we are interested in is the study of the paths of the vertices and of the center of gravity of the triangle. First, however, we will express the results already obtained as q-series.

Since we know the network of periods is rectangular, let us choose a rectangle standing upon a smaller side.

Put $2\omega_1 = \pi$ † and then $2\omega_2 = ir\pi$ where $r > 1$. Hence we have $q = e^{i\pi}\omega_2/\omega_1 = e^{-r\pi}$, and the larger we take r the smaller q becomes.

pu expressed as a q-series is ‡

$$
\begin{aligned} pu = &-(\eta_1/\omega_1) + (\pi/2\omega_1)^2 1/\sin^2(\pi u/2\omega_1) \\ &- 2(\pi/\omega_1)^2 \sum_{n=1}^{\infty} [nq^{2n}/(1-q^{2n})] \cos nu(\pi/\omega_1). \end{aligned}
$$

For $2\omega_1 = \pi$, we have

$$
pu = -2\eta_1/\pi + 1/\sin^2 u - 8 \sum [nq^{2n}/(1-q^{2n})] \cos 2nu;
$$

and for $\tau = \pi/3$, we obtain

$$
p\tau = -2\eta_1/\pi + 1/\sin^2 \pi/3 - 8 \sum [nq^{2n}/(1-q^{2n})] \cos 2n(\pi/3).
$$

we are interested in the values for pu obtained for u moving along a line from ω_2 to $\omega_2 + 2\pi$ or, what is the same thing, for $u = \omega_2 + v$.

---

* Jacobi, *Fundamenta Nova*; Halphen, *Traité des Fonctions Elliptiques*, Vol. 1, p. 425.

† When we have put $2\omega_1 = \pi$, our unit is fixed and we are talking about a particular triangle. To generalize we merely multiply x, y, and z by μ (an arbitrary constant), and the discussion is the same.

‡ Halphen, *Traité des Fonctions Elliptiques*, Vol. 1, p. 426.

$$p(\omega_2 + v) = -2\eta_1/\pi - 8\sum[nq^n/(1-q^{2n})]\cos 2nv.^*$$

Expressed as q-series, we have

5.1
$$\begin{cases} x = 1/(p\tau - pu) = (3/4)[1 - 6q\cos 2v + q^2(15 + 6\cos 4v) + \cdots]. \\ y = 1/[p\tau - p(u+\tau)] = (3/4)[1 - 6q\cos 2(v + \pi/3) \\ \qquad + q^2(15 + 6\cos 4(v+\pi/3) + \cdots]. \\ z = 1/[p\tau - p(u-\tau)] = (3/4)[1 - 6q\cos 2(v - \pi/3) \\ \qquad + q^2(15 + 6\cos 4(v-\pi/3) + \cdots]. \end{cases}$$

5.2 $$k_1 = x + y + z = (9/4)(1 + q^2 + \cdots).$$

5.3 $$k_3 = xyz = (27/64)(1 + 18q^2 + \cdots).$$

5.4 $$g_2 = 4/3 + 320(q^2 + \cdots).^\dagger$$

5.5 $$g_3 = 8/27 - (2^6\cdot 7/3)(q^2 + \cdots).^\dagger$$

5.6 $$\Delta = 2^{12}q^2 + \cdots.^\dagger$$

We are now prepared to explain our choice of a rectangular cell standing upon a smaller side. If $q = 0$ we see from 5.6 that the discriminant is zero. But this says $k_1^3 = 27k_3$ or that $a = b = c$ which is the equilateral case. Once equilateral the triangle stays equilateral, and the vertices move on a circle. It is easy to show that the triangle will never become equilateral unless it is that way initially. We shall consider the nearly equilateral case, hence we want q to be small. We can also express Λ, θ, ϕ, and ψ as q-series. We had $d\Lambda/dv = abc/2k_1k_3^{1/2}$; hence, $d\Lambda/dv = (2/3^{1/2})(1 + 57q^2/4 + \cdots)$.

5.7 $$\therefore\quad \Lambda = (2/3^{1/2})(1 + 57q^2/4 + \cdots)v + \Lambda_0,$$

where $\Lambda_0 = 0$. Also

$$d\theta/dv = (1/k_1^{1/2})\{1 + 2/[k_1(p\tau - pu) - 1]\},$$

$$d\theta/dv = (2/3)[2 - 9q\cos 2v - 3q^2/2 + 45q^2\cos 4v/2 + \cdots].$$

5.8 $$\therefore\quad \theta = 4v/3 - 3q\sin 2v - q^2v + 15q^2\sin 4v/4 + \cdots + \theta_0.$$

We can choose our triangle to make $\theta_0 = 0$. In our original conditions when $v = 0$, $u = \omega_2$ and our triangle is isosceles. Let us choose our base line to be initially parallel to the side a. We must determine the constants of integration for ϕ and ψ.

$$\phi - k = (4/3)v - 3q\sin 2(v + \pi/3) \\ - q^2(v + \pi/3) + (15/4)q^2\sin 4(v + \pi/3) + \cdots,$$

$$\phi_0 - k = -3q\sin(2\pi/3) - q^2\cdot\pi/3 + (15/4)q^2\sin(4\pi/3) + \cdots,$$

whence

5.9 $$\phi - \phi_0 = (4/3)v + 3q[\sin 2\pi/3 - \sin 2(v + \pi/3)] \\ - q^2v + (15/4[\sin 4(v + \pi/3) - \sin(4\pi/3)] + \cdots.$$

* Halphen, *Traité des Fonctions Elliptiques*, Vol. 1, p. 426.

† Harkness and Morley, *Theory of Functions*, pp. 322-324.

In the same way

5.10 $$\psi - \psi_0 = (4/3)v - 3q[\sin 2\pi/3 + \sin 2(v - \pi/3) - q^2 v + (15/4)[\sin 4(v - \pi/3) + \sin (4\pi/3)] + \cdots.$$

In order to determine ϕ_0 and ψ_0 consider

$$a = k_1 - 1/(p\tau - pu) = 3/2 - (9/2)q \cos 2v + [45/2 - (9/2) \cos 4v]\, q^2 + \cdots.$$

For $v = 0$; $a = a_0$,

$$\therefore \quad a_0 = (3/2)[1 - 3q + 12q^2 + \cdots],$$

and in a similar way, we find

$$b_0 = c_0 = (3/2)[1 + 3q/2 + 33q^2/2 + \cdots].$$

Hence

5.11 $$\cos \phi_0 = -\frac{a_0}{2b_0} = -\frac{1 - 3q + 12q^2 + \cdots}{2 + 3q + 33q^2 + \cdots}$$

From this we get

5.12 $$e^{i\phi_0} + e^{-i\phi_0} = -1 + 9q/2 - 9q^2/4 + \cdots.$$

$$\therefore \quad e^{i\phi_0} = \omega + (i3^{3/2}/2)\omega^2 q + \cdots,$$

and $$i\phi_0 = \log (\omega + (i3^{3/2}/2)\omega^2 q + \cdots.$$

To evaluate ψ_0, we note that it is equal to $-\phi_0$, and hence $\cos \psi_0 = \cos \phi_0$. When we solve 5.12 for $e^{i\phi_0}$, we have two roots resulting from a quadratic equation. They represent the values of $e^{i\phi_0}$ and $e^{i\psi_0}$ respectively. Hence

$$e^{i\psi_0} = \omega^2 - (i3^{3/2}/2)\omega q + \cdots,$$

and $$i\psi_0 = \log [\omega^2 - i3^{3/2}/2)\omega q + \cdots].$$

Formula 5.11 is very important in that it fixes the value of q once the initial lengths of the sides of the triangle are given.

6. *The Paths of the Vertices and Centroid.* First consider the path of A. $e^{i\theta}$ is the turn from the base line to the side a. A is moving along some path with a direction $e^{(\pi+\theta)i} = -e^{i\theta}$. Since this equals $dA/d\Lambda$, we have

$$dA/d\Lambda = \exp\{i[4v/3 - 3q \sin 2v - q^2 v + 15q^2 \sin 4v/2 + \cdots]\}$$
$$= -\exp(4iv/3)\{1 - 3iq \sin 2v + q^2[-(9/4) - iv + (9/4)\cos 4v + (15i/4)\sin 4v] + \cdots\}, \quad \text{and}$$

$$d\Lambda/dv = 2(1 + 57q^2/4 + \cdots)/3^{1/2}.$$

$$\therefore \quad dA/dv = -(2/3^{1/2}) \exp[i(4/3)v]\,[1 - 3iq \sin 2v$$
$$+ q^2(12 - iv + 9\cos 4v/4 + 15i \sin 4v/4) + \cdots]$$
$$= (-2/3^{1/2}) \exp[i(4/3)v]\{1 - 3q[\exp(2iv) - \exp(-2iv)]/2$$
$$+ q^2[48 - 4iv + 12\exp(4iv) - 3\exp(-4iv)]/4 + \cdots\}.$$

Put $\exp(2iv) = t^3$, then $dv = 3dt/2it$, and

$$dA/dt = i3^{1/2}[t - 3q(t^4 - t^{-2})/2$$
$$+ (q^2/4)(48t - 6t \log t + 12t^7 - 3t^{-5}) + \cdots].$$

6.1 $$\therefore \quad A = A_0 + i3^{1/2}\{t^2/2 - (3q/10t)(t^6 + 5)$$
$$+ (q^2/16)[102t^2 - 12t^2 \log t + 6t^8 + 3t^{-4}] + \cdots\}.$$

We can determine A_0 by taking $A = 0$ when $t = 1$. This represents the path along which A moves. The logarithmic term tells us the path is not closed but continually shifts over the plane.

By a similar method the equations of the paths of B and C, expanded as far as the first degree term in q, are found to be

6.2 $$B = B_0 + i3^{1/2}\left(\frac{\omega t^2}{2} - \frac{i3^{3/2}qt^2}{4} - \frac{3\omega q(\omega t^6 + 5\omega^2)}{10t} + \cdots\right),$$

6.3 $$C = C_0 + i3^{1/2}\left(\frac{\omega^2 t^2}{2} + \frac{i3^{3/2}qt^2}{4} - \frac{3\omega^2 q(\omega^2 t^6 + 5\omega)}{10t} + \cdots\right)$$

where B_0 and C_0 are determined in the same manner as A_0.

To obtain the path of the centroid g we have, noting that $3g = A+B+C$,

that $$g = (A_0 + B_0 + C_0)/3 + 3^{3/2}q/2it + \cdots,$$
where $$A_0 + B_0 + C_0 = 3^{3/2}iq/2 + \cdots.$$
Hence $$g = (3^{3/2}q/2i)[1/(t-1)] + \cdots.$$

Periodic Orbits in the Problem of Three Bodies with Repulsive and Attractive Forces.

By Daniel Buchanan.

1. *Introduction.* This paper deals with periodic orbits described by two mutually repellant infinitesimal bodies which are attracted by a finite body. The forces of repulsion and attraction are assumed to vary according to the Newtonian law of the inverse square. Two types of periodic orbits for this system were obtained by Rawles.* In the first type, which will be here designated as the *circular orbits,* the repellant particles move in equal circles the planes of which are parallel. The line joining the centres of these circles is normal to their planes and is bisected by the centre of gravity of the finite body. The particles remain on the same generating line of the cylinder through these circles.

In the orbits of the second type, here designated as the *arc orbits,* the three bodies remain in the same plane. The infinitesimal bodies oscillate in arcs of curves, which are symmetrically situated with respect to the finite body. Langmuir † first calculated these orbits by numerical integration and they are also discussed by Van Vleck.‡

The problem considered in the present paper deals with periodic oscillations in the vicinity of the circular orbits. Only the construction of these orbits is made but the convergence of the solutions obtained is assured by a theorem due to MacMillan.§ The author begs to acknowledge the assistance of Mr. H. D. Smith, M. A.,¶ in checking certain algebraic expressions in the construction and in making the computation for the numerical examples.

Second genus orbits in the vicinity of the arc orbits have also been obtained by the author but they are discussed in another article.‖

---

* Rawles, "Two Classes of Periodic Orbits with Repelling Forces," *Bulletin of the American Mathematical Society*, Vol. 34, No. 5 (1928), pp. 618-630.

† Langmuir, *Physical Review*, Vol. 17 (1921), pp. 339-353.

‡ Van Vleck, "Quantum Principles and Line Spectra," *Bulletin of the National Research Council*, Vol. 10, Part 4, No. 54, p. 89.

§ MacMillan, *Transactions of the American Mathematical Society*, Vol. 13, No. 2, pp. 146-158.

¶ Smith, A thesis submitted in the Department of Mathematics for the degree of M. A. in the University of British Columbia.

‖ Buchanan, "Second Genus Orbits for the Helium Atom," *Transactions of the Royal Society of Canada*, Third Series, Vol. 23, Sec. 3 (1929), pp. 227-245.

As there is a similarity between the three bodies in this problem and the helium atom, we shall refer to the finite body as the nucleus and to the particles as electrons. No use, however, is made of the quantum mechanics nor of Larmor's theorem.*

2. *The Circular Orbits.* The units of time and space will be chosen so that the gravitational constant of attraction is unity. Let k^2 denote the ratio of the repulsion to the attraction. Then the force function of the system is

$$U = 1/\rho_1 + 1/\rho_2 - k^2/\Delta,$$

where ρ_1 and ρ_2 are the distances between the electrons and the nucleus, and Δ is the distance between the electrons. If we take a system of rectangular coördinates with the origin at the nucleus and denote the coördinates of the electrons as (x_j, y_j, z_j), $(j = 1, 2)$, then the differential equations defining their motion are

$$\begin{aligned} &x_j'' = \partial U/\partial x_j, \quad y_j'' = \partial U/\partial y_j, \quad z_j'' = \partial U/\partial z_j, \\ (1) \qquad &\rho_j^2 = x_j^2 + y_j^2 + z_j^2, \qquad\qquad (j = 1, 2), \\ &\Delta^2 = (x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2. \end{aligned}$$

When the restrictions

$$(2) \qquad x_1 = -x_2, \quad y_1 = y_2, \quad z_1 = z_2$$

are made, as in Rawles' paper, the differential equations become

$$(3) \qquad \begin{aligned} x'' &= -x/\rho^3 + k^2/4x^2, \\ y'' &= -y/\rho^3, \\ z'' &= -z/\rho^3, \end{aligned}$$

where the subscripts 1 or 2 have been dropped. These equations possess the integrals

$$(4) \qquad \begin{aligned} \tfrac{1}{2}(x'^2 + y'^2 + z'^2) &= 1/\rho - k^2/4x + \text{const.}, \\ y'z - yz' &= \text{const.} \end{aligned}$$

The solutions of the differential equations are

$$(5) \qquad \begin{aligned} x &= (k^2/4)^{1/3} = m, \text{ say}, \\ y &= (1 - m^2)^{1/2} \sin (t - t_0), \\ z &= (1 - m^2)^{1/2} \cos (t - t_0), \end{aligned}$$

which are the circular solutions obtained by Rawles. They denote the circles

* Larmor, *Philosophical Magazine*, V, Vol. 44 (1897), p. 503; Richardson, *The Electron Theory of Matter* (1916), p. 258.

with centres at $(\pm m, 0, 0)$, radii $(1-m^2)^{1/2}$ and whose planes are parallel to the yz-plane. The electrons rotate in these orbits from the positive z-axis to the positive y-axis. If the solutions are to be real, m^2 cannot exceed unity. When $m^2 = 1$, however, the solutions reduce to point circles but this simple case will be excluded from our consideration.

We shall refer only to the one circle, viz., that having its centre at $(m, 0, 0)$.

Orbits of Three Dimensions.

3. *The Differential Equations.* Let the motion be referred to a system of rotating axes x, η, ξ. The x-axis remains unchanged while the $\eta\xi$-axes rotate in the yz-plane in the direction in which the electrons move and with their angular velocity. Further, let $\eta = -y$, $\xi = z$ at $t = t_0$. Then the necessary transformations are

$$(6) \qquad \begin{aligned} y &= -\eta \cos (t - t_0) + \xi \sin (t - t_0), \\ z &= \eta \sin (t - t_0) + \xi \cos (t - t_0), \end{aligned}$$

and the differential equations of motion (3) become

$$(7) \qquad \begin{aligned} x'' &= -x/\rho^3 + k^2/4x^2, \\ \eta'' + 2\xi' - \eta &= -\eta/\rho^3, \\ \xi'' - 2\eta - \xi &= -\xi/\rho^3. \end{aligned}$$

A particular solution of these equations is

$$(8) \qquad x = m, \quad \eta = 0, \quad \xi = (1 - m^2)^{1/2},$$

which are the equations of the circular orbit with respect to the rotating axes.

In order to determine deviations from the circular orbit, let

$$(9) \qquad \begin{aligned} x &= m + \gamma p, \\ \eta &= 0 + \gamma q, \\ \xi &= (1 - m^2)^{1/2} + \gamma r, \\ t - t_0 &= (1 + \delta)^{1/2} \tau, \end{aligned}$$

where

p, q, r are new dependent variables,
γ is a parameter representing the scale factor of the new orbits,
δ is a constant depending upon γ,
τ is the new independent variable.

When equations (9) are substituted in (7) and the factor γ is divided out,

the following differential equations are found, the dots denoting derivation with respect to τ;

$$
\begin{aligned}
&\ddot{p} + 3(1+\delta)(1-m^2)p - 3(1+\delta)m(1-m^2)^{1/2}r \\
&\qquad = (1+\delta)[\gamma P_2 + \gamma^2 P_3 + \cdots + \gamma^j P_{j+1} + \cdots], \\
(10)\quad &\ddot{q} + 2(1+\delta)^{1/2}\dot{r} = (1+\delta)[\gamma Q_2 + \cdots + \gamma^j Q_{j+1} + \cdots], \\
&\ddot{r} - 2(1+\delta)^{1/2}\dot{q} - 3(1+\delta)(1-m^2)r - 3(1+\delta)m(1-m^2)^{1/2}p \\
&\qquad = (1+\delta)[\gamma R_2 + \cdots + \gamma^j R_{j+1} + \cdots],
\end{aligned}
$$

where P_j, Q_j, R_j $(j = 2, 3, \cdots)$ are polynomials in p, q, r of degree j. In P_j and R_j, q enters to even degrees only, while in Q_j it enters to odd degrees only. So far as the computation has been carried out we have

$$
\begin{aligned}
P_2 &= 3(1/m + 3m/2 - 5m^3/2)p^2 + 3mq^2/2 \\
&\quad - 3m(2 - 5m^2/2)r^2 + 3(1-m^2)^{1/2}(1-5m^2)pr, \\
P_3 &= (3/2 - 15m^2 + 35m^4/2)p^3 + (15m/2)(1-m^2)^{1/2}(7m^2-3)p^2r \\
&\quad + (3/2)(1-5m^2)pq^2 - 3(2 - 35m^2/2 + 35m^4/2)pr^2 \\
&\quad - (15m/2)(1-m^2)^{1/2}q^2r + 5(2 - 7m^2/2)r^3, \\
Q_2 &= 3mpq + 3(1-m^2)^{1/2}qr, \\
Q_3 &= (3/2)(1-5m^2)p^2q - 15m(1-m^2)^{1/2}pqr + 3q^3/2 \\
&\quad - 3(2 - 5m^2/2)r^2q, \\
R_2 &= (3/2)(1-m^2)^{1/2}(1-5m^2)p^2 - 3m(4-5m^2)pr \\
&\quad + (3/2)(1-m^2)^{1/2}q^2 - 3(1-m^2)^{1/2}(1-5m^2/2)r^2, \\
R_3 &= (5m/2)(1-m^2)^{1/2}(7m^2-3)p^3 \\
&\quad + 3[1/2 + 15m^2 - 35m^4/2 - (5m/2)(1-m^2)^{1/2}]p^2r \\
&\quad - (15m/2)(1-m^2)^{1/2}pq^2 + 15m(1-m^2)^{1/2}(2 - 7m^2/2)pr^2 \\
&\quad + 3[1/2 - 5m(1-m^2)^{1/2}]q^2r \\
&\quad - [27/2 - 15m^2 - (35/2)(1-m^2)^{5/2}]r^3.
\end{aligned}
$$

On integrating (10, b) we obtain

$$
(11)\qquad \dot{q} = -2(1+\delta)^{1/2}r + C + (1+\delta)\int(\gamma Q_2 + \cdots + \gamma^j Q_{j+1}\cdots)d\tau
$$

where C is the constant of integration. As q and r are later developed as power series in γ we shall put

$$
(12)\qquad C = C_1^{(0)} + C_1^{(1)}\gamma + \cdots + C_1^{(n)}\gamma^n + \cdots.
$$

When the substitutions are made for C in (11) and for $\dot{q}$ in (10, c) we obtain, on repeating (10, a) and (11) for reference,

$$
\begin{aligned}
(13)\qquad & \ddot{p} + 3(1+\delta)(1-m^2)p - 3(1+\delta)m(1-m^2)^{1/2}r \\
& \quad = (1+\delta) \sum_{j=1}^{\infty} \gamma^j P_{j+1}, \\
& \dot{q} = -2(1+\delta)^{1/2}r + (1+\delta)\int \sum_{j=1}^{\infty} \gamma^j Q_{j+1}d\tau + \sum_{j=0}^{\infty} C_1^{(j)}\gamma^j, \\
& \ddot{r} + (1+\delta)(1+3m^2)r - 3(1+\delta)m(1-m^2)^{1/2}p \\
& \quad = (1+\delta) \sum_{j=1}^{\infty} \gamma^j R_{j+1} + 2(1+\delta)^{1/2} \sum_{j=0}^{\infty} C_1^{(j)}\gamma^j \\
& \quad + 2(1+\delta)^{3/2} \int \sum_{j=1}^{\infty} \gamma^j Q_{j+1} d\tau.
\end{aligned}
$$

We shall now take (13) as the three defining equations for p, q, r.

4. *The Equations of Variation and their Solutions.* If we consider only the terms of the equations (13) which are independent of γ we obtain the equations of variation. They are

$$
(14)\qquad
\begin{aligned}
& \ddot{p} + 3(1-m^2)p - 3m(1-m^2)^{1/2}r = 0, \\
& \dot{q} + 2r = C_1^{(0)}, \\
& \ddot{r} + (1+3m^2)r - 3m(1-m^2)^{1/2}p = 2C_1^{(0)}.
\end{aligned}
$$

The first and third equations of (14) are independent of the second and will be considered first. We shall make use of the operator D to denote $d/d\tau$. Then (14, a) and (14, c) may be expressed as

$$
(15)\qquad
\begin{aligned}
& [D^2 + 3(1-m^2)]\, p - 3m(1-m^2)^{1/2}r = 0, \\
& -3m(1-m^2)^{1/2}p + [D^2 + 1 + 3m^2]r = 2C_1^{(0)}.
\end{aligned}
$$

The functional determinant of these equations is

$$
(16)\qquad
\mathfrak{D} = \begin{vmatrix} D^2 + 3(1-m^2), & -3m(1-m^2)^{1/2} \\ -3m(1-m^2)^{1/2}, & D^2 + 1 + 3m^2 \end{vmatrix}
= D^4 + 4D^2 + 3(1-m^2).
$$

On equating $\mathfrak{D}$ to zero, as in the method of solving sets of linear differential equations with constant coefficients, we find the roots

$$D^2 = -2 + (1+3m^2)^{1/2}, \quad -2 - (1+3m^2)^{1/2}.$$

As m^2 must be less than 1 in order that the circular solutions shall be real, both roots for D^2 are therefore negative. If we put

$$-2 + (1+3m^2)^{1/2} = -\sigma_1^2, \quad -2 - (1+3m^2)^{1/2} = -\sigma_2^2,$$

then

$$D = \pm i\sigma_1, \quad \pm i\sigma_2,$$

and the complementary functions of (15) are thus found to be

$$(17)\qquad \begin{aligned} p &= A_1 e^{i\sigma_1\tau} + A_2 e^{-i_1\sigma\tau} + A_3 e^{i\sigma_2\tau} + A_4 e^{-i\sigma_2\tau}, \\ r &= B_1 e^{i\sigma_1\tau} + B_2 e^{-i\sigma_1\tau} + B_3 e^{i\sigma_2\tau} + B_4 e^{-i\sigma_2\tau}, \end{aligned}$$

where A_j, B_j, $(j = 1, \cdots, 4)$ are constants of integration. Only four of these constants are independent as the following relations hold,

$$(18)\qquad \begin{aligned} A_j &= \omega_\nu B_j, \qquad\qquad (j = 1, 2, 3, 4;\ \nu = 1, 2), \\ \omega_1 &= 3m(1 - m^2)^{1/2}/[1 - 3m^2 + (1 + 3m^2)^{1/2}], \\ \omega_2 &= 3m(1 - m^2)^{1/2}/[1 - 3m^2 - (1 + 3m^2)^{1/2}]. \end{aligned}$$

There are therefore three sets of generating solutions, viz.,

$$\text{I}\qquad \begin{aligned} p &= \omega_1(B_1 e^{i\sigma_1\tau} + B_2 e^{-i\sigma_1\tau}), \\ r &= B_1 e^{i\sigma_1\tau} + B_2 e^{-i\sigma_1\tau}; \\ &\text{Period} = P_1 = 2\pi/\sigma_1. \end{aligned}$$

$$\text{II}\qquad \begin{aligned} p &= \omega_2(B_3 e^{i\sigma_2\tau} + B_4 e^{-i\sigma_2\tau}), \\ r &= B_3 e^{i\sigma_2\tau} + B_4 e^{-i\sigma_2\tau}; \\ &\text{Period} = P_2 = 2\pi/\sigma_2. \end{aligned}$$

$$\text{III}\qquad \begin{aligned} p &= \omega_1(B_1 e^{i\sigma_1\tau} + B_2 e^{-i\sigma_1\tau}) + \omega_2(B_3 e^{i\sigma_2\tau} + B_4 e^{-i\sigma_2\tau}), \\ r &= B_1 e^{i\sigma_1\tau} + B_2 e^{-i\sigma_1\tau} + B_3 e^{i\sigma_2\tau} + B_4 e^{-i\sigma_2\tau}; \\ &\text{Period} = P_3 = n_2 P_1 = n_1 P_2. \end{aligned}$$

The last solutions, III, exist only when σ_1 and σ_2 are commensurable, i. e., when

$$\sigma_1/\sigma_2 = n_1/n_2,$$

where n_1 and n_2 are relatively prime integers.

Orbits are constructed in the sequel by using only the first two generating solutions. The construction of orbits having generating solutions III was attempted but abandoned on account of the complexity of the problem.

5. *Outline of the Construction of Periodic Solutions.* There is the same construction for orbits having the generating solutions I or II except for the subscripts 1 and 2, respectively, on σ and ω. We shall therefore drop these subscripts and restore them in the final solutions.

We propose to show that p, q, r, δ can be determined as power series in γ so that p, q, r shall be periodic with the period $P (= P_1$ or $P_2)$ and shall satisfy certain initial conditions, to be discussed presently. Accordingly we put

$$(19)\qquad \begin{aligned} p &= \sum_{j=0}^{\infty} p_j\gamma^j, & q &= \sum_{j=0}^{\infty} q_j\gamma^j, \\ r &= \sum_{j=0}^{\infty} r_j q^j, & \delta &= \sum_{j=1}^{\infty} \delta_j\gamma^j. \end{aligned}$$

Let these substitutions be made in (13) and let the resulting equations be cited as (13′). On equating the coefficients of the various powers of γ in (13′) we obtain sets of differential equations in p_j, q_j, r_j. We propose to show that these equations can be integrated and that the various δ_j and the constants of integration at each step can be determined so that p_j, q_j, r_j shall be periodic and shall satisfy the initial conditions, now to be discussed.

6. *The Initial Conditions.* It will be observed in the next section that at each step of the integration four arbitrary constants arise which are not determined by the periodicity conditions. We therefore impose four initial conditions. Let us suppose that

$$\dot{p}(0)=\dot{r}(0)=q(0)=0, \qquad r(0)\neq 0.$$

As r carries the factor γ in (9) we may take $r(0)=1$ without loss of generality. When these initial conditions are imposed upon (19) we obtain

$$\begin{aligned} &\dot{p}_j(0)=\dot{r}_j(0)=q_j(0)=0, && (j=0,1,2\cdots), \\ &r_1(0)=1, \quad r_j(0)=0, && (j=2,3,4,\cdots). \end{aligned} \tag{20}$$

7. *Construction of the Solutions.*

Terms independent of γ. When we equate the coefficients of the terms in (13′) which are independent of γ we obtain equations which are the same as (14) except for the subscript 0 on p, q and r. The solutions which have the period P_1 or P_2, except for certain terms in τ, are

$$\begin{aligned} p_0 &= \omega(B_1^{(0)}e^{i\sigma\tau}+B_2^{(0)}e^{-i\sigma\tau})+2m(1-m^2)^{-1/2}C_1^{(0)}, \\ q_0 &= (2i/\sigma)(B_1^{(0)}e^{i\sigma\tau}-B_2^{(0)}e^{-i\sigma\tau})-3C_1^{(0)}\tau+C_2^{(0)}, \\ r_0 &= B_1^{(0)}e^{i\sigma\tau}+B_2^{(0)}e^{-i\sigma\tau}+2C_1^{(0)}, \end{aligned} \tag{21}$$

where B and C, here and henceforth, with various subscripts and superscripts are constants of integration.

In order to satisfy the periodicity conditions we must put $C_1^{(0)}=0$. When we impose the condition $\dot{p}_0(0)=0$ we obtain $B_1^{(0)}=B_2^{(0)}$, and consequently the condition $\dot{r}_0(0)=0$ is satisfied. Then from $q_0(0)=0$ we obtain $C_2^{(0)}=0$ and from $r_0(0)=1$ we have

$$B_1^{(0)}=B_2^{(0)}=1/2.$$

The periodic solutions at this step which satisfy the initial conditions then become

$$p_0=\omega\cos\sigma\tau, \quad q_0=-(2/\sigma)\sin\sigma\tau, \quad r_0=\cos\sigma\tau. \tag{22}$$

Terms in γ. The differential equations arising from the terms in γ in (13′) are

$$
(23)\quad
\begin{aligned}
&[D^2+3(1-m^2)]p_1-3m(1-m^2)^{1/2}r_1=P^{(1)},\\
&-3m((1-m^2)^{1/2}p_1+[D^2+1+3m^2]r_1=R^{(1)}+2C_1^{(1)},\\
&\dot{q}_1=-2r_1+C_1^{(1)}+\int Q^{(1)}d\tau,
\end{aligned}
$$

where

$$
\begin{aligned}
P^{(1)}&=a_0^{(1)}+\delta_1a_1^{(1)}\cos\sigma\tau+a_2^{(1)}\cos2\sigma\tau,\\
Q^{(1)}&=\qquad\delta_1b_1^{(1)}\sin\sigma\tau+b_2^{(1)}\sin2\sigma\tau,\\
R^{(1)}&=c_0^{(1)}+\delta_1c_1^{(1)}\cos\sigma\tau+c_2^{(1)}\cos2\sigma\tau;\\
a_0^{(1)}&=(3/2)(1/m+3m/2-5m^3/2)\omega^2+(3/2)(1-m^2)^{1/2}(1-5m^2)\\
&\quad-3m(1-1/\sigma^2-5m^2/4),\\
a_1^{(1)}&=-3(1-m^2)\omega+3m(1-m^2)^{1/2},\\
a_2^{(1)}&=(3/2)(1/m-3m/2-5m^3/2)\omega^2+(3/2)(1-m^2)^{1/2}(1-5m^2)\\
&\quad-3m(1+1/\sigma^2-5m^2/4),\\
b_1^{(1)}&=1,\quad b_2^{(1)}=-(3/\sigma)[m\omega+(1-m^2)^{1/2}],\\
c_0^{(1)}&=(3/4)(1-m^2)^{1/2}(1-5m^2)\omega^2-3m(2-5m^2/2)\omega\\
&\quad-3(1-m^2)^{1/2}(1/2-1/\sigma^2-5m^2/4),\\
c_1^{(1)}&=3m(1-m^2)^{1/2}+(1-3m^2),\\
c_2^{(1)}&=(3/4)(1-m^2)^{1/2}(1-5m^2)\omega^2-3m(2-5m^2/2)\omega\\
&\quad-(3/2)(1-m^2)^{1/2}(1+1/\sigma^2-5m^2/2).
\end{aligned}
$$

The solutions of (23, a and b) will be considered first as (23, c) depends upon r_1. The complementary functions of (23 a) and (23 b) are

$$
(24)\quad
\begin{aligned}
p_1&=\omega(B_1^{(1)}e^{i\sigma\tau}+B_2^{(1)}e^{-i\sigma\tau})+2m(1-m^2)^{-1/2}C_1^{(1)},\\
r_1&=B_1^{(1)}e^{i\sigma\tau}+B_2^{(1)}e^{-i\sigma\tau}+2C_1^{(1)}.
\end{aligned}
$$

The particular integrals of p_1 and r_1, expressed symbolically, are

$$
(25)\quad
\begin{aligned}
p_1&=\frac{[D^2+1+3m^2]P^{(1)}+3m(1-m^2)^{1/2}R^{(1)}}{D^4+4D^2+3(1-m^2)},\\
r_1&=\frac{3m(1-m^2)^{1/2}P^{(1)}+[D^2+3(1-m^2)]R^{(1)}}{D^4+4D^2+3(1-m^2)}.
\end{aligned}
$$

In order that p_1 and r_1 shall be periodic the coefficients of $\cos\sigma\tau$ in the numerators of the above expressions must vanish, inasmuch as $-\sigma^2$ is a root of the denominators. Hence

$$
(26)\quad
\begin{aligned}
&\delta_1[a_1^{(1)}\{-\sigma^2+1+3m^2\}+c_1^{(1)}\{3m(1-m^2)^{1/2}\}]=0,\\
&\delta_1[a_1^{(1)}\{3m(1-m^2)^{1/2}\}+c^{(1)}\{-\sigma^2+3(1-m^2)\}]=0.
\end{aligned}
$$

The functional determinant of $\delta_1a_1^{(1)}$ and $\delta_1c_1^{(1)}$ in the above equations is

$$\sigma^4-4\sigma^2+3(1-m^2),$$

and this vanishes as $-\sigma^2$ is a root of $\mathfrak{D}$ in (16). Therefore the two equations in (26) are equivalent. They are satisfied only by $\delta_1 = 0$. The particular integrals then become

$$(27)\qquad \begin{aligned} p_1 &= \alpha_0^{(1)} + \alpha_2^{(1)} \cos 2\sigma\tau, \\ r_1 &= \gamma_0^{(1)} + \gamma_2^{(1)} \cos 2\sigma\tau, \end{aligned}$$

where

$$\alpha_0^{(1)} = \frac{1+3m^2}{3(1-m^2)} a_0^{(1)} + \frac{m}{(1-m^2)^{1/2}} c_0^{(1)},$$

$$\alpha_2^{(1)} = \frac{(1-4\sigma^2+3m^2)a_2^{(1)} + 3m(1-m^2)^{1/2}c_2^{(1)}}{16\sigma^4 - 16\sigma^2 + 3(1-m^2)},$$

$$\gamma_0^{(1)} = \frac{m}{(1-m^2)^{1/2}} a_0^{(1)} + c_0^{(1)},$$

$$\gamma_2^{(1)} = \frac{3m(1-m^2)^{1/2}a_2^{(1)} - \{4\sigma^2 - 3(1-m^2)\}c_2^{(1)}}{16\sigma^4 - 16\sigma^2 + 3(1-m^2)}.$$

When (24) and (27) are combined we obtain the complete solutions for p_1 and r_1.

The third equation of (23) can now be integrated, the integral being

$$q_1 = (2i/\sigma)(B_1^{(1)}e^{i\sigma\tau} - B_2^{(1)}e^{-i\sigma\tau}) - (3c_1^{(1)} + 2\gamma_0^{(1)})\tau + \beta_2^{(1)} \sin 2\sigma\tau + C_2^{(1)},$$

where

$$\beta_2^{(1)} = (3/4\sigma^3)[m\omega + (1-m^2)^{1/2} - 2\gamma_2^{(1)}].$$

On applying the periodicity and initial conditions to the complete solutions for p_1, q_1, r_1 we obtain

$$C_1^{(1)} = -(2/3)\gamma_0^{(1)}, C_2^{(1)} = 0,$$
$$B_1^{(1)} = B_2^{(1)} = (1/6)\gamma_0^{(1)} - (1/2)\gamma_2^{(1)}.$$

The desired solutions at this step are thus found to be

$$(28)\qquad \begin{aligned} p_1 &= F_0^{(1)} + F_1^{(1)} \cos \sigma\tau + F_2^{(1)} \cos 2\sigma\tau, \\ q_1 &= G_1^{(1)} \sin \sigma\tau + G_2^{(1)} \sin 2\sigma\tau, \\ r_1 &= H_0^{(1)} + H_1^{(1)} \cos \sigma\tau + H_2^{(1)} \cos 2\sigma\tau, \end{aligned}$$

where

$$\begin{aligned} &F_0^{(1)} = \alpha_0^{(1)} - (4m/3)(1-m^2)^{-1/2}\gamma_0^{(1)}, \\ &F_1^{(1)} = 2B_1^{(1)}\omega,\ F_2^{(1)} = \alpha_2^{(1)}, \\ &G_1^{(1)} = -(4/\sigma)B_1^{(1)},\ G_2^{(1)} = \beta_2^{(1)}, \\ &H_0^{(1)} = -(1/3)\gamma_0^{(1)},\quad H_1^{(1)} = 2B_1^{(1)},\ H_2^{(1)} = \gamma_2^{(1)}. \end{aligned}$$

Terms in γ^2. It will be necessary to consider the terms in γ^2 in (13′) before the induction to the general term can be made. These terms are

$$
(29)\quad
\begin{aligned}
&[D^2+3(1-m^2)]\,p_2-3m(1-m^2)^{1/2}r_2=P^{(2)},\\
&-3m(1-m^2)^{1/2}p_2+(D^2+1+3m^2)r_2=R^{(2)}+2C_1^{(2)},\\
&\dot q_2=-2r_2+\int Q^{(2)}d\tau-\delta_2 r_0+C_1^{(2)},
\end{aligned}
$$

where

$$
\begin{aligned}
P^{(2)}&=a_0^{(2)}+(\delta_2 d_1^{(2)}+a_1^{(2)})\cos\sigma\tau\\
&\quad+a_2^{(2)}\cos 2\sigma\tau+a_3^{(2)}\cos 3\sigma\tau,\\
R^{(2)}&=c_0^{(2)}+(\delta_2 d_2^{(2)}+c_1^{(2)})\cos\sigma\tau\\
&\quad+c_2^{(2)}\cos 2\sigma\tau+c_3^{(2)}\cos 3\sigma\tau,\\
Q^{(2)}&=b_1^{(2)}\sin\sigma\tau+b_2^{(2)}\sin 2\sigma\tau+b_3^{(2)}\sin 3\sigma\tau,\\
d_1^{(2)}&=3m(1-m^2)^{1/2}-3(1-m^2)\omega,\\
d_2^{(2)}&=1-3m^2+3m(1-m^2)^{1/2}\omega.
\end{aligned}
$$

The values of the various a's, b's, and c's were computed by Mr. Smith but his results are omitted here.

The complementary functions and the particular integrals of the first two equations of (29) are the same as (24) and (25), respectively, with the appropriate changes in subscripts and superscripts. The equations similar to (26) which must be satisfied in order that the particular integrals for p_1 and r_1 shall be periodic, are

$$
(30)\quad
\begin{aligned}
&(1-\sigma^2+3m^2)[\delta_2 d_1^{(2)}+a_1^{(2)}]+3m(1-m^2)^{1/2}[\delta_2 d_2^{(2)}+c_1^{(2)}]=0,\\
&3m(1-m^2)^{1/2}[\delta_2 d_1^{(2)}+a_1^{(2)}]+\{-\sigma^2+3(1-m^2)\}[\delta_2 d_2^{(2)}+c_1^{(2)}]=0.
\end{aligned}
$$

The determinant of the coefficients of the expressions in the brackets [ ] is the same here as in (26) and therefore vanishes. Hence the above equations are identical and can be satisfied by a proper choice of the single arbitrary δ_2. The required value of δ_2 is

$$
(31)\quad \delta_2=\frac{(\sigma^2-1-3m^2)a_1^{(2)}-3m(1-m^2)^{1/2}c_1^{(2)}}{(1-\sigma^2+3m^2)d_1^{(2)}+3m(1-m^2)^{1/2}d_2^{(2)}}.
$$

When δ_2 is thus determined, the complete solutions for p_2 and r_2 will be periodic and will have the form

$$
(32)\quad
\begin{aligned}
p_2&=\omega(B_1^{(2)}e^{i\sigma\tau}+B_2^{(2)}e^{-i\sigma\tau})+2m(1-m^2)^{-1/2}C_1^{(2)}\\
&\quad+\alpha_0^{(2)}+\alpha_2^{(2)}\cos 2\sigma\tau+\alpha_3^{(2)}\cos 3\sigma\tau,\\
r_2&=B_1^{(2)}e^{i\sigma\tau}+B_2^{(2)}e^{-i\sigma\tau}+2C_1^{(2)}\\
&\quad+\gamma_0^{(2)}+\gamma_2^{(2)}\cos 2\sigma\tau+\gamma_3^{(2)}\cos 3\sigma\tau,
\end{aligned}
$$

where the α's and γ's are linear in the a's and c's.

On substituting (32) in (29 c) and integrating we obtain

$$
\begin{aligned}
q_2&=(2i/\sigma)(B_1^{(2)}e^{i\sigma\tau}-B_2^{(2)}e^{-i\sigma\tau})+(3C_1^{(2)}+2\gamma_0^{(2)})\tau\\
&\quad+C_2^{(2)}+\beta_1^{(2)}\sin\sigma\tau+\beta_2^{(2)}\sin 2\sigma\tau+\beta_3^{(2)}\sin 3\sigma\tau,
\end{aligned}
$$

where

$$\beta_1^{(2)} = -(1/\sigma)\delta_2 - (1/\sigma^2)b_1^{(2)},$$
$$\beta_2^{(2)} = -(1/\sigma)\gamma_2^{(2)} - (1/4\sigma^2)b_2^{(2)},$$
$$\beta_3^{(2)} = -(2/3\sigma)\gamma_3^{(2)} - (1/9\sigma^2)b_3^{(2)}.$$

When the periodicity and initial conditions are applied we have

$$C_1^{(2)} = -(2/3)\gamma_0^{(2)}, \quad C_2^{(2)} = 0,$$
$$B_1^{(2)} = B_2^{(2)} = (1/6)\gamma_0^{(2)} - \gamma_2^{(2)} + \gamma_3^{(2)}.$$

The solutions at the third step are therefore

$$p_2 = \sum_{\nu=0}^{3} F_\nu^{(2)} \cos \nu\sigma\tau,$$

$$q_2 = \sum_{\nu=1}^{3} G_\nu^{(3)} \sin \nu\sigma\tau,$$

$$r_2 = \sum_{\nu=0}^{3} H_\nu^{(2)} \cos \nu\sigma\tau,$$

where

$$F_0^{(2)} = 2m(1-m^2)^{-\frac{1}{2}}c_1^{(2)} + \alpha_0^{(2)},$$
$$F_1^{(2)} = 2\omega B_1^{(2)} \quad F_j^{(2)} = \alpha_j^{(2)}, \qquad (j = 2, 3,),$$
$$G_1^{(2)} = -(4/\sigma)B_1^{(2)} + \beta_1^{(2)}, G_j^{(2)} = \beta_j^{(2)}, \qquad (j = 2, 3),$$
$$H_0^{(2)} = 2c_1^{(2)} + \gamma_0^{(2)},$$
$$H_1^{(2)} = 2B_1^{(2)}, H_j^{(2)} = \gamma_j^{(2)}, \qquad (j = 2, 3).$$

8. *Induction to the General Term.* Let us suppose that the p_j, q_j, r_j have all been determined for $j = 0, \cdots, n-1$ and that they are of the form

$$
\begin{aligned}
p_j &= \sum_{\nu=0}^{j+1} F_\nu^{(j)} \cos \nu\sigma\tau, \\
q_j &= \sum_{\nu=1}^{j+1} G_\nu^{(j)} \sin \nu\sigma\tau, \\
r_j &= \sum_{\nu=0}^{j+1} H_\nu^{(j)} \cos \nu\sigma\tau, \qquad (j = 0, \cdots, n-1),
\end{aligned}
\tag{33}
$$

where the $F_\nu^{(j)}$, $G_\nu^{(j)}$, $H_\nu^{(j)}$ are functions of m. Further, let us suppose that $\delta_1, \cdots, \delta_{n-1}$ have been uniquely determined. We wish to show from these assumptions, from the differential equations, and from the initial and periodicity conditions that p_n, q_n, r_n have the same form as (33) for $j = n$, and that δ_n is a uniquely determined constant.

The differential equations obtained by equating the coefficients of γ^n in (13′) are

$$
\begin{aligned}
&[D^2 + 3(1-m^2)]p_n - 3m(1-m^2)^{\frac{1}{2}}r_n = P^{(n)}, \\
&-3m(1-m^2)^{\frac{1}{2}}p_n + [D^2 + 1 + 3m^2]r_n = R^{(n)} + 2C_1^{(n)}, \\
&\dot{q}_n = -2rn + \int Q^{(n)}d\tau + C_1^{(n)} - \delta_n r_0,
\end{aligned}
\tag{34}
$$

where

$$\begin{aligned} P^{(n)} &= -3\delta_n(1-m^2)p_0 + 3\delta_n m(1-m^2)^{1/2}r_0 \\ &\quad + \text{terms in } p_j,\ q_j,\ r_j,\ \delta_j, \\ R^{(n)} &= 3\delta_n m(1-m^2)p_0 - \delta_n(1+3m^2)r_0 \\ &\quad + \text{terms in } p_j,\ q_j,\ r_j,\ \delta_j, \\ Q^{(n)} &= \text{terms in } p_j,\ q_j,\ r_j,\ \delta_j, \qquad (j=0,\cdots,n-1;\ \delta_0=0). \end{aligned}$$

The undetermined constant δ_n enters the right members only where it is expressed and not in the other terms. In $P^{(n)}$ and $R^{(n)}$ the powers of the q's are even while in $Q^{(n)}$ they are odd. Hence $P^{(n)}$ and $R^{(n)}$ are sums of cosines of multiples of $\sigma\tau$ while $Q^{(n)}$ is a sum of sines of multiples of $\sigma\tau$. They have the form

$$\begin{aligned} P^{(n)} &= a_0^{(n)} + (d_1^{(n)}\delta_n + a_1^{(n)})\cos\sigma\tau + \cdots + a^{(n)}_{n+1}\cos(n+1)\sigma\tau, \\ R^{(n)} &= c_0^{(n)} + (d_2^{(n)}\delta_n + c_1^{(n)})\cos\sigma\tau + \cdots + c^{(n)}_{n+1}\cos(n+1)\sigma\tau, \\ Q^{(n)} &= b_1^{(n)}\sin\sigma\tau + \cdots + b^{(n)}_{n+1}\sin(n+1)\sigma\tau, \end{aligned}$$

The complementary functions of (34, a and b) and the terms arising from $2C_1^{(2)}$ in (34 b) are

$$\begin{aligned} p_n &= \omega(B_1^{(n)}e^{i\sigma\tau} + B_2^{(n)}e^{-i\sigma\tau}) + 2m(1-m^2)^{-1/2}C_1^{(n)}, \\ r_n &= B_1^{(n)}e^{i\sigma\tau} + B_2^{(n)}e^{-i\sigma\tau} + 2C_1^{(n)}. \end{aligned}$$

The symbolic expressions for the particular integrals are the same as (25) with the appropriate changes in subscripts and superscripts. As at the previous steps the coefficients of $\cos\sigma\tau$ in the numerators of these expressions must vanish in order that p_n and q_n shall be periodic. We thus arrive at the two equations

$$\begin{aligned} (1-\sigma^2+3m^2)(d_1^{(n)}\delta_n + a_1^{(n)}) + 3m(1-m^2)^{1/2}(d_2^{(n)}\delta_n + c_1^{(n)}) &= 0, \\ 3m(1-m^2)^{1/2}(d_1^{(n)}\delta_n + a_1^{(n)}) + [-\sigma^2 + 3(1-m^2)](d_2^{(n)}\delta_n + c_1^{(n)}) &= 0. \end{aligned}$$

Since the functional determinant in these equations vanishes, the two equations are equivalent and can be satisfied by solving either for δ_n. Thus

$$\delta_n = \frac{-(-\sigma^2+1+3m^2)a_1^{(n)} - 3m(1-m^2)^{1/2}c_1^{(n)}}{(-\sigma^2+1+3m^2)d_1^{(n)} + 3m(1-m^2)^{1/2}d_2^{(n)}}.$$

With this choice of δ_n the particular integrals will be periodic and will have the form

$$\begin{aligned} p_n &= \alpha_0^{(n)} + \alpha_2^{(n)}\cos 2\sigma\tau + \cdots + \alpha^{(n)}_{n+1}\cos(n+1)\sigma\tau, \\ r_n &= \gamma_0^{(n)} + \gamma_2^{(n)}\cos 2\sigma\tau + \cdots + \gamma^{(n)}_{n+1}\cos(n+1)\sigma\tau, \end{aligned}$$

On substituting the complete solution for r_n in (34 c) and integrating we obtain

$$q_n = (2i/\sigma)(B_1^{(n)}e^{i\sigma\tau} - B_2^{(n)}e^{-i\sigma\tau}) - (3C_1^{(n)} + 2\gamma_0^{(n)})\tau$$
$$+ C_2^{(n)} + \sum_{\nu=2}^{n+1} \beta_\nu^{(n)} \sin \nu\sigma\tau,$$

and in order that this solution shall be periodic we must put

$$C_1^{(n)} = -(2/3)\gamma_0^{(n)}.$$

When the initial conditions are applied we obtain

$$C_2^{(n)} = 0, \quad B_1^{(n)} = B_2^{(n)} = \text{a constant.}$$

Hence p_n, q_n and r_n have the same form as (33) when $j = n$. This completes the induction. The construction of the solutions can therefore be carried on to any desired degree of accuracy.

The two sets of solutions can be obtained by restoring the subscripts 1 or 2 to ω and σ.

9. *The Final Form of the Solutions.* On substituting the various values for p_j, q_j, r_j in (19) and the results in (9) we obtain

$$x = m + \sum_{j=0}^{\infty} \left(\sum_{\nu=0}^{j+1} F_\nu^{(j)} \cos \nu\sigma\tau \right) \gamma^{j+1},$$
$$\eta = 0 + \sum_{j=0}^{\infty} \left(\sum_{\nu=1}^{j+1} G_\nu^{(j)} \sin \nu\sigma\tau \right) \gamma^{j+1},$$
$$\xi = (1 - m^2)^{1/2} + \sum_{j=0}^{\infty} \left(\sum_{\nu=0}^{j+1} H_\nu^{(j)} \cos \nu\sigma\tau \right) \gamma^{j+1},$$
$$\tau = (1 + \sum_{j=1}^{\infty} \delta_j \gamma^j)^{-1/2} (t - t_0).$$

In the above equations m, γ and t_0 are the only parameters which remain arbitrary; m denoting the scale factor of the circular orbits, γ that of the periodic oscillations near these orbits, and t_0 the epoch. By substituting for η and ξ in the equations

$$y = -\eta \cos (t - t_0) + \xi \sin (t - t_0),$$
$$z = \eta \sin (t - t_0) + \xi \cos (t - t_0),$$

we may obtain the corresponding values of y and z. There are two sets of values, x_1, y_1, z_1; x_2, y_2, z_2, corresponding to the two electrons, but they are not independent inasmuch as the restrictions (2) hold

$$x_1 = -x_2, \quad y_1 = y_2, \quad z_1 = z_2.$$

10. *Numerical Example.* Mr. Smith assigned the values

$$k^2 = .5, \quad m = .5, \quad \gamma = .05, \quad t_0 = 0,$$

and on completing the integrations up to the terms in p_2, q_2 and r_2 he obtained

$$p = -.0025 + .064 \cos \sigma\tau + .025 \cos 2\sigma\tau - .002 \cos 3\sigma\tau,$$
$$q = -.19 \sin \sigma\tau + .03 \sin 2\sigma\tau - .0003 \sin 3\sigma\tau,$$
$$r = .0043 + .077 \cos \sigma\tau - .017 \cos 2\sigma\tau - .00027 \cos 3\sigma\tau.$$

Using the subscript 1 on ω and σ he found

$$\sigma_1 = .825, \quad P_1 = 12\pi/5, \quad \text{nearly.}$$

Values of t were then taken at approximately 30° intervals as t ranges from 0° to 2160°, that is, through the complete period, and the numerical values of x_1, y_1, and z_1 were computed. The values obtained near the beginning and near the end of the period are found in the accompanying Table.

t^0	x_1	y_1	z_1
0	.560	.00	.95
30	.555	.54	.77
60	.536	.87	.33
90	.515	.87	—.19
120	.488	.64	—.59
150	.464	.26	—.78
180	.450	—.10	—.80
210	.440	—.43	—.66
240	.475	—.65	—.46
270	.458	—.81	—.14
300	.480	—.81	.26
330	.506	—.60	.65
360	.530	—.17	.83
. . .	. . .	. . .	. . .
. . .	. . .	. . .	. . .
1800	.530	.17	.93
1836	.500	.68	.57
1890	.458	.81	—.14
1926	.443	.60	—.51
1980	.448	.10	—.80
2016	.470	—.36	—.76
2070	.515	—.87	—.19
2106	.542	—.81	.43
2160	.560	0	.95

A check was made on the work by making use of the *vis viva* integral (4 a). Various sets of computed values for x_1, y_1, z_1 and their derivatives were used and the constant in the *vis viva* integral was found to range from 2.17 to 2.31.

The accompanying diagrams give the projections of the oscillations on the coördinate planes. The circular orbit is not shown in Fig. 2. Its projections in Fig. 1 and Fig. 3 are the y- and z-axes respectively.

11. *Two-Dimensional Orbits.* Two-dimensional periodic oscillations near the circular orbits can be readily found by neglecting the terms in x in the preceding construction. These orbits are coplanar with the circular orbits.

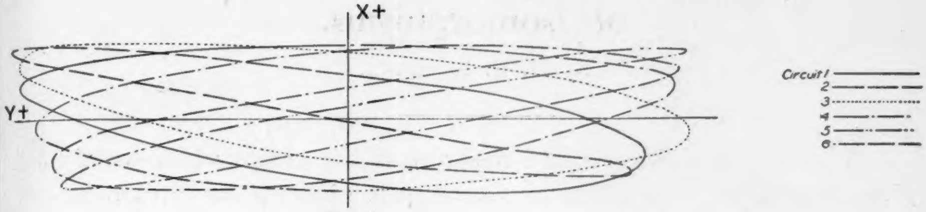


Fig. 1

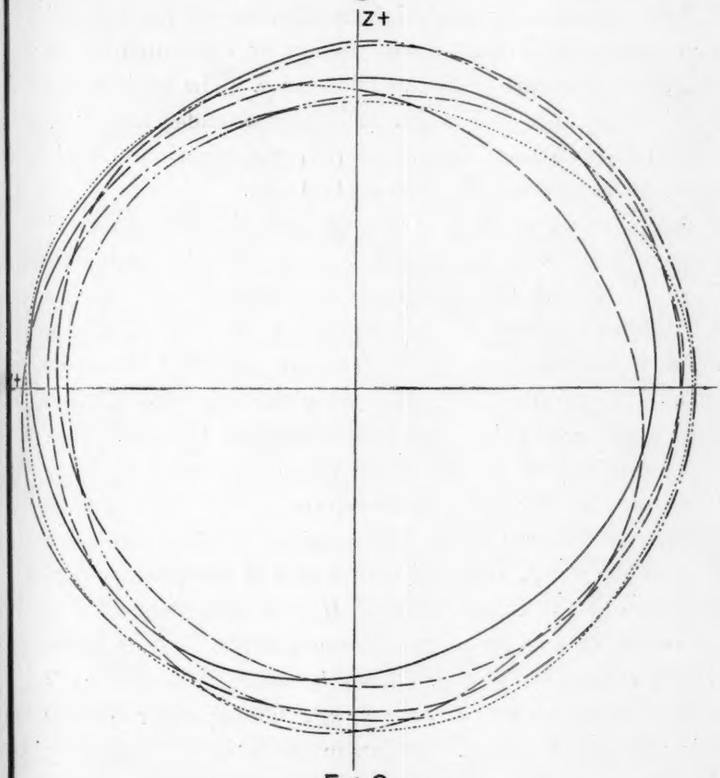


Fig. 2

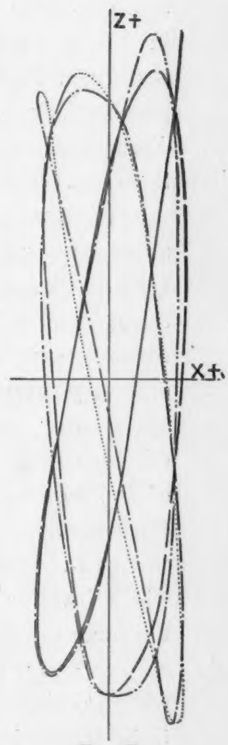


Fig. 3

The actual construction was carried out but as no peculiarities were found it is omitted. Mr. Smith computed an orbit and found curves similar to those in Fig. 2.

THE UNIVERSITY OF BRITISH COLUMBIA,
VANCOUVER, CANADA.

On the Groups Which Contain a Given Invariant Subgroup and Transform It According to a Given Operator in Its Group of Isomorphisms.

By H. R. BRAHANA.

A method by which one may construct all the groups which contain a given group H as an invariant subgroup of prime index p was given recently by Professor Miller.* In the papers cited the method was applied and several theorems were introduced which accomplished simplifications of the method in special cases, mostly cases in which H was abelian or the isomorphism performed on H by an operator outside H was of order p. The subject was presented by Professor Miller to a class which the writer attended and after discussion it was decided to investigate the possible wider application of these theorems. The results of this investigation are offered here.

We consider a group H and a group G of order $p \cdot h$ which contains H as an invariant subgroup of prime index p. Let t_1 be an operator outside H. Its p-th power will be in H, and G is generated by t_1 and H. Following the method used by Miller (*loc. cit.*) we may consider G to be written as a regular group in which H is intransitive but is transitive on the h letters of each of p constituents. The operator t_1 permutes these constituents cyclically. Let t be an operator on the $p \cdot h$ letters which permutes the transitive constituents of H in the same way as t_1, but which transforms every operator of H into itself. Then the operator t_1t^{-1} will transform each of the transitive constituents into itself and will transform the operators of H in the same way as t_1. Let $t_1t^{-1} = s_1's_2' \cdots s_p'$, where s_i' is that part of the product t_1t^{-1} which involves only letters of the i-th constituent H_i. s_1' transforms H_1 in the same way as some operator s_1 in its group of isomorphisms. Let us define $s_2, s_3, \cdots, s_p$ by the relation $t^{-1}s_it = s_{i+1}$. Then t_1 which is $s_1's_2' \cdots s_p't$ performs the same transformation on H as $s_1s_2 \cdots s_pt$. The operator $Q = t_1^p$ is in H and hence is permutable with t. Q transforms H in the same way as $s_1^ps_2^p \cdots s_p^p$. The operator $Q \cdot s_1^{-p}s_2^{-p} \cdots s_p^{-p}$ which we shall denote by $\bar{s}_0'\bar{s}_0'' \cdots \bar{s}_0^{(p)}$, where $\bar{s}_0^{(i)}$ is that part of the product which involves only

* (1) *Proceedings of the National Academy of Sciences*, Vol. 14 (1928), p. 819. See also (2) *loc. cit.*, p. 918; and (3) *Transactions of the American Mathematical Society*, Vol. 2 (1901), p. 264, and (4) *American Journal of Mathematics*, Vol. 24 (1902), p. 395, in which he described and used the method in the construction of prime power groups.

letters of the i-th transitive constituent, is permutable with every operator of H and also with t. This operator is in the conjoint of H and moreover it is transformed into itself by s_i since this is true of both Q and s_j. Now let us consider the operator $U = \bar{s}_0' s_1 s_2 \cdots s_p t$. U transforms H in the same manner as t_1 and its p-th power is $\bar{s}_0' \bar{s}_0'' \cdots \bar{s}_0^{(p)} s_1^p s_2^p \cdots s_p^p$ which is Q. Therefore, $\{H, U\}$ is simply isomorphic with $\{H, t_1\}$.

Conversely, if there exists an operator s in the group of isomorphisms of H whose p-th power is an inner isomorphism and an operator Q in H which transforms the operators of H in the same way as s^p and is invariant under s, then the operator $\bar{s}_0'$ and consequently the operator U and the group G exist. Therefore,

A necessary and sufficient condition that there exists a group G of order $p \cdot h$ in which the operators of a given invariant subgroup H are transformed according to an operator s in its group of isomorphisms whose p-th power is an inner isomorphism is that there exists an operator Q of H which transforms the operators of H according to s^p and is invariant under s.

The operator $\bar{s}_0'$, and consequently U also, is completely determined by Q and s. s does not determine Q completely but determines it as one of a set of operators of H each of which transforms H in the same manner as s^p and each of which is permutable with s. The operators of H which transform H in the same manner as s^p may all be obtained from one of them by multiplying it in turn by operators from the central of H. The operators of the central of H which are permutable with s form a subgroup C which when s is not identity is the central of G and does not depend on Q. Therefore,

Every group G which contains a given group H invariantly as a subgroup of prime index p and transforms it according to a given operator s, not identity, in its group of isomorphisms contains a central C which depends only on H and s.

The order of s is of necessity a multiple of p, but in any group G the operator U may be so chosen that it transforms H according to an operator s whose order is a power of p, for if the order of the transformation performed by U is $m \cdot p^a$ where m is prime to p then U^m will transform H according to an operator s whose order is a power of p. The groups $\{H, U\}$ and $\{H, U^m\}$ are evidently the same. We shall therefore assume in what follows that the order of s is a power of p.

A necessary and sufficient condition that for a given H and s there exist a group G is given in the first theorem. That such a group need not always exist was shown by Professor Miller.* We shall accordingly in what follows

---

* *loc. cit.*, (1) p. 821.

assume that one such group exists for the H and s under consideration and investigate the question of the existence of other groups.

If the given group is $\{H, U\}$ where $U^p = Q$, every possible group determined by H and s is generated by H and an operator which transforms H according to s and which has $C_i \cdot Q$ for a p-th power, where C_i is some operator of C. Since s is of order p^a, $U^{p^a} = R \cdot Q'$, where both R and Q' are in C, the order of R is prime to p, and the order of Q' is a power of p. Therefore, the group $\{H, U\}$ will contain an operator $\bar{U} = R^k \cdot U$ which transforms the operators of H according to s and whose p^a-th power is Q' of order a power of p. Since the groups $\{H, U\}$ and $\{H, \bar{U}\}$ are the same, we may assume that the order of Q is a power of p.

Now any other group that corresponds to H and s may be obtained by taking H and s_0U where s_0 is chosen so that $s_0s_0' \cdots s_0^{(p)}$ is an operator in C; and though every operator of C will give an s_0 and every s_0 determines a group, it follows from the preceding paragraph that the number of distinct groups cannot exceed the order of the Sylow subgroup of order p^γ in C.

If R is any operator of C then $(RU)^p = R^pQ$. Therefore, the group obtained by taking s_0 to correspond to R^p is the same as that obtained by taking s_0 to be identity. We have the theorem:

The number of groups which contain a given group H invariantly as a subgroup of index p and transform its operators according to a given operator s in its group of isomorphisms is not more than one greater than the number of operators which are not p-th powers in the Sylow subgroup of order p^γ in the central C.

An operator of G which transforms H in the same manner as U must be the product of U and an operator from the central of H, and if it has for a p-th power the product of Q and an operator from the Sylow subgroup $C_p{}^\gamma$ of order p^γ of C the operator from the central of H must be from its Sylow subgroup $H_p{}^\beta$ of order p^β. Let R be such an operator, let $U^{-1}RU = R_1R$, and let $U^{-1}R_iU = R_{i+1}R_i$. Then since $U^p = Q$, we have $U^{-p}RU^p = R_pR_{p-1}^{\binom{p}{1}} R_{p-2}^{\binom{p}{2}} \cdots R_1 \quad R = R$, where the exponents are the binomial coefficients. From this we get

$$R_pR_{p-1}^{\binom{p}{1}} R_{p-2}^{\binom{p}{2}} \cdots R_1^{\binom{p}{1}} = 1. \tag{1}$$

Then $(RU)^p = U^p \cdot R_pR_{p-1}^{\binom{p+1}{1}} \cdots R_1^{\binom{p+1}{2}} \cdot R^p$, which in view of (1) becomes

$$(RU)^p = Q \cdot R_{p-1}R_{p-2}^{\binom{p}{1}} R_{p-3}^{\binom{p}{2}} \cdots R_1^{\binom{p}{2}} \cdot R^p. \tag{2}$$

If R' is another operator in the central of H the operator $(R'U)^p$ will be the same as the right member of (2) where the R_i is replaced by R_i'. Then $(R'RU)^p = Q \cdot (R'_{p-1}R_{p-1}) \; (R'_{p-2}R_{p-2})^{\binom{p}{1}} \cdots (R_1'R_1)^{\binom{p}{2}} (R'R)^p$ $= Q \cdot R'_{p-1}R'^{\binom{p}{1}}_{p-2} \cdots R'^{\binom{p}{2}}_1 R'^p \cdot R_{p-1}R^{\binom{p}{1}}_{p-2} \cdots R^{\binom{p}{2}}_1 R^p$. Hence the operators of $C_p{}^\gamma$ which with Q determine p-th powers of operators of G which transform H in the same manner as U form a group; we shall denote this group by C_M. Moreover, the set of operators $R_{p-1}R^{\binom{p}{1}}_{p-2} R^{\binom{p}{2}}_{p-3} \cdots R^{\binom{p}{2}}_1 R^p$, where R is allowed to go through a set of independent generators of the Sylow subgroup of order p^β of the central of H generate a group which contains every operator in the central of H which can be written in that form. The cross-cut of this group and C is C_M.

If $C_p{}^\gamma$ is arranged in co-sets with respect to C_M, a choice of s_0 which makes the product of Q and one operator of a particular co-set the p-th power of an operator which transforms H in the same manner as U, makes the product of Q and every operator of that co-set such a p-th power. Therefore,

The number of groups determined by a given H and s does not exceed the order of the quotient group of $C_p{}^\gamma$ with respect to C_M.

It is true that we may determine an s_0 for each operator of the quotient group of $C_p{}^\gamma$ with respect to C_M and that each such s_0 determines a group G which has a new set of operators for p-th powers of operators which transform H in the same way as U, but we may not conclude therefrom that there are that many distinct groups G, due to the possibility of isomorphisms of H which are permutable with s. This will become more apparent when we consider certain restrictions on H and s.

The method of procedure indicated in the proof of the preceding theorem is quite readily carried out when both p and the number of invariants of $C_p{}^\gamma$ are small. Often, however, the result may be arrived at indirectly in a simpler manner. From the form of the right member of (2) we notice that every operator of C_M is in the group H_M generated by the p-th powers of operators in the Sylow subgroup of order p^β in the central of H and the $(p-1)$-th derived group of this Sylow subgroup with respect to U. Since C_M is in C, C_M will be in the cross-cut of H_M and C; we shall denote this cross-cut by C_L.

We shall now consider some of the subgroups of C_M. In any case where we can show that such a subgroup coincides with C_L, we may conclude that C_M coincides with this subgroup.

Let us consider an operator R in the central of H whose p-th power is in C_p^{γ}. Then $U^{-1}R^pU = R_1^pR^p$ which must be R^p. Therefore, R_1 and each of the succeeding R_i's must be of order p or 1. Then from (1) we see that R_p must be identity, which requires R_{p-1} to be in C_p^{γ}. Moreover, (2) reduces to $(RU)^p = Q \cdot R_{p-1}R^p$. If R_{p-1} is identity then the operator R^p is in C_M.* The R's for which the corresponding R_{p-1}'s are identity form a group and their p-th powers form a group which is in C_M and which we shall denote by C_I.

If one of the operators R_{p-1} above is the p-th power of an operator S in C_p^{γ}, then $(S^{-1}RU)^p = Q \cdot S^{-p}R_{p-1}R^p = Q \cdot R^p$. Then R^p is in C_M. The product of two such R's fulfills the same conditions, as do the operators R which determine C_I. Thus we have determined a group C_J which is contained in C_M and contains C_I.

The $(p-1)$-th derived group with respect to U of the set of operators of the central of H whose p-th powers are in C_p^{γ} is, as we have seen, contained in C_p^{γ} and is of type $1, 1, \cdots$. The group C_I contains all of those R_{p-1}'s which are p-th powers in C_p^{γ}. For each of the independent generators of the group of R_{p-1}'s which are not in C_I we may determine an operator $R_{p-1}R^p$, any one of which is obtained from a given one by multiplying the latter by some operator from C_I. The group C_K determined by these operators and C_J is contained in C_M and contains C_I.

These three groups may be described as follows: C_I is composed of the set of p-th powers of the set of operators in the central of H whose p-th powers are in C_p^{γ} and whose $(p-1)$-th commutators are identity; C_J is obtained by removing the restriction that the $(p-1)$-th commutators be identity and requiring that they be p-th powers in C_p^{γ}; and C_K is obtained by extending C_J by means of a definite operator for each of the remaining generators of the $(p-1)$-th derived group of the set of operators in the central of H whose p-th powers are in C_p^{γ}.

If we suppose that R is an operator in the group H_p^{β} for which R_{p-1} and $R_{p-1}R_{p-2}^{\binom{p}{1}} \cdots R_1^{\binom{p}{2}} R^p$ are in C_p^{γ}, we note first that R_{p-1} is of order p, since $U^{-p}R_{p-2}U^p = R_{p-1}^{\binom{p}{1}} R^1{}_{p-2} = R_{p-2}$. Then from $U^{-p}R_{p-3}U^p = R_{p-1}^{\binom{p}{2}} R_{p-2}^{\binom{p}{1}} R_{p-3} = R_{p-3}$ it follows that R_{p-2} is also of order p. By repetition of this process we may show that every R_i is of order p, and that therefore $R_{p-1}R_{p-2}^{\binom{p}{1}} \cdots R_1^{\binom{p}{2}} R^p$ becomes $R_{p-1}R^p$. Hence under the conditions on R its p-th power must be in

* This includes two of Miller's theorems: (1) R is in C, *loc. cit.* (1), p. 820; and (2) R_1 is in C, is of order p, *loc. cit.* (3), p. 265.

C_p^γ and the subgroup C_K of C_M cannot be extended by an operator corresponding to an R_{p-1} which is invariant.

To continue to a consideration of the R_{p-1}'s which are non-invariant would be to give a complete determination of C_M for which a method has already been pointed out. From the foregoing a number of conclusions concerning special cases may be drawn; we shall give three.

(a) *If the* $(p-1)$*-th derived group of* H_p^β *is identity, then* C_M *coincides with* C_I*; if it is composed of p-th powers in* C_p^γ*, then* C_M *coincides with* C_J*; if it is composed of invariant operators, then* C_M *coincides with* C_K*.*

(b) *If the group of p-th powers of operators of* H_p^β *is contained in* C_p^γ*, then the* $(p-1)$*-th derived group of* H_p^β *is in* C_p^γ *and* C_M *coincides with* C_K*.*

(c) *If* C_p^γ *coincides with the group of p-th powers of operators of* H_p^β*, and if the* $(p-1)$*-th derived group of* H_p^β *is contained in the group of p-th powers of operators of* C_p^γ*, then* C_M *coincides with* C_J *and with* C_p^γ*, and therefore there is but one group corresponding to* H *and* s*.*

If R is an operator of H_p^β which is transformed into its k-th power by U, then $(RU)^p = Q \cdot R^{1+k+k^2+\cdots+k^{p-1}}$. If this operator $R^{1+k+\cdots+k^{p-1}}$ is in C_p^γ it is in C_M.* If we have determined C_M this gives us no new information, but if we are determining C_I, C_J, or C_K it gives additional information concerning C_M whenever $R^{1+k+\cdots+k^{p-1}}$ is not in C_K.

Thus far we have placed no restrictions on H or s. Let us now suppose that s is of order p. The operator s^p is permutable with every operator of H. Since H always contains at least one operator, namely identity, which is invariant under H and s, Q always exists. Therefore,

For a given H *and an* s *of order* p *in its group of isomorphisms there exists at least one group* G *which contains* H *invariantly and transforms it according to* s*.*

If H is abelian s must be of order p. This makes no change in the procedure in the determination of C_M since that depended only on the operators of H_p^β, which were permutable with each other and with s^p. However, when H is abelian it is the direct product of its Sylow subgroups and its group of isomorphisms is the direct product of the groups of isomorphisms of its Sylow

* **This theorem is given by Miller for H abelian, *loc. cit.* (2), p. 918.**

subgroups. These Sylow subgroups are abelian and therefore the group of isomorphisms of H contains invariant operators which perform α-automorphisms * on the Sylow subgroups. Hence when H is abelian the group obtained by extending H by means of s_0U is simply isomorphic with that obtained by extending H by means of s_0^kU where k is prime to p, and therefore in the determination of s_0 it is necessary to consider but one operator, and that of highest order, from any cyclic subgroup. Hence,

If H is abelian the number of groups G which contain H invariantly and transform it according to a given operator of order p in its group of isomorphisms is not more than one greater than the number of cyclic groups which are not contained in cyclic groups of higher order of the quotient group of C_p^γ with respect to C_M.

If H is cyclic all the above groups are cyclic and therefore there cannot be more than two groups for a given s.† If H is cyclic and s does not leave invariant the operators of highest order of H_p^γ, then C_M coincides with C_p^γ except that when $p=2$ and C_2^γ is of order 2 then we have $R_1R^2=1$ where R is the operator of order 4. Thus when H is cyclic there are two groups, only if s leaves the operators of highest order of H_p^β invariant, or, when $p=2$, transforms them into their inverses.

The theorem just stated for H abelian is not true when H is non-abelian and we shall conclude by giving an example to prove it.

Let H be $\{s_1, s_2, s_3\}$ where s_1, s_2, and s_3 satisfy the conditions

$$(1)\qquad \begin{cases} s_1^p=1, & s_2^p=1, \quad s_1s_2=s_2s_1, \\ s_3^p=s_1, & s_3^{-1}s_2s_3=s_1s_2. \end{cases}$$

It is obvious that H is a non-abelian group of order p^3 and contains an abelian subgroup of order p^2 and type 1, 1. Now let us consider the groups G and $\bar{G}$ obtained by adjoining operators s_4 and $\bar{s}_4$ which satisfy respectively the relations

$$(2)\qquad s_4^{-1}s_3s_4=s_2s_3, \quad s_4s_2=s_2s_4, \quad s_4^p=s_1,$$

and

$$(\bar{2})\qquad \bar{s}_4^{-1}s_3\bar{s}_4=s_2s_3, \quad \bar{s}_4s_2=s_2\bar{s}_4, \quad \bar{s}_4^p=s_1^2.$$

The groups G and $\bar{G}$ are both of order p^4, the operators s_4 and $\bar{s}_4$ perform the same transformation on operators of H and are of the same order. In

---

* Burnside, *Theory of Groups* (1911), p. 113. See also, Miller, *loc. cit.* (2), p. 266.

† Cf. Miller, *loc. cit.* (2), p. 919.

the one case, however, $s_3^p = s_4^p$ and in the other $s_3^{2p} = \bar{s}_4^p$. We shall prove that the two groups are in general not simply isomorphic.

Each group contains an abelian subgroup of order p^3, $\{s_2, s_4\}$ and $\{s_2, \bar{s}_4\}$ respectively. These subgroups are each invariant and we proceed to show that in any simple isomorphism of G and $\bar{G}$ they must correspond. The commutator subgroups K and $\bar{K}$ in each case is $\{s_1, s_2\}$ and these two groups must correspond. Every operator in G may be written $k_i s_4^k s_3^l$ where k_i is in K and $l < p$. Since s_3 is not permutable with s_2 an abelian group of order p^3 in G must be such that each operator can be written in the form $k_i s_4^k$. Hence, the group $\{s_2, s_4\}$ is the only abelian group of order p^3 in G and $\{s_2, \bar{s}_4\}$ is the only abelian group of order p^3 in $\bar{G}$. Therefore, in any simple isomorphism of G and $\bar{G}$, $A = \{s_2, s_4\}$ and $\bar{A} = \{s_2, \bar{s}_4\}$ must correspond.

If we attempt to set up the simple isomorphism we must select operators σ_2, σ_4, and σ_3, the first two in A and the third in G outside of A which satisfy the same relations as s_2, $\bar{s}_4$, and s_3 of $\bar{G}$. Looking first to the transformations of operators in A and $\bar{A}$ by G and $\bar{G}$ we observe that every operator of G which is not in A transforms the operators of A in the same way as some power of s_3. $\bar{G}$ is generated by s_2, $\bar{s}_4$, and s_3 which satisfy the relations:

$$(4)\qquad \begin{cases} s_3^{-1}s_2s_3 = s_1s_2 = (\bar{s}_4^p)^{(p+1)/2}\cdot s_2 = \bar{s}_4^{p(p+1)/2}\cdot s_2, \\ s_3^{-1}\bar{s}_4s_3 = s_2^{-1}\bar{s}_4^{-p(p+1)/2}\cdot \bar{s}_4. \end{cases}$$

Every operator of A may be written $s_2^a s_4^b$. Then if G and $\bar{G}$ are simply isomorphic there exist operators $\sigma_2 = s_2^d s_4^e$, $\sigma_4 = s_2^a s_4^b$, and $\sigma_3 = s_3^c$ which satisfy relations (4) when substituted for s_2, $\bar{s}_4$, and s_3 respectively. Operators s_2, s_4 and s_3 satisfy relations

$$(5)\qquad \begin{cases} s_3^{-1}s_2s_3 = s_1s_2 = s_4^p s_2, \\ s_3^{-1}s_4s_3 = s_2^{-1}s_1^{-1}s_4 = s^{-1}s_4^{-p}\cdot s_4. \end{cases}$$

From these we get

$$(6)\qquad \begin{cases} s_3^{-c}s_2s_3^c = s_1^c s_2, \text{ and } s_3^{-c}s_2^a s_3^c = s_1^{ac}s_2^a, \\ s_3^{-c}s_4s_3^c = s_1^{-c(c+1)/2}s_2^{-c}\cdot s_4, \text{ and } s_3^{-c}s_4^b s_3^c = s_1^{-bc(c+1)/2}s_2^{-bc}s_4^b. \end{cases}$$

Combining these we get

$$(7)\qquad \begin{cases} s_3^{-c}s_2^a s_4^b s_3^c = s_2^{-bc}s_1^{ac-bc(c+1)/2}\cdot s_2^a s_4^b = s_2^{-bc}s_4^{acp-bcp(c+1)/2}\cdot s_2^a s_4^b, \\ \text{and a similar one where } d \text{ and } e \text{ are put in place of } a \text{ and } b. \end{cases}$$

If in (4) we substitute σ_2, σ_4, and σ_3 and compare the right members with the corresponding right members of (7) we get

$$(8) \qquad \begin{cases} ec \equiv 0, \bmod p, \\ dc \equiv b(p+1)/2, \bmod p, \\ bc - d \equiv 0, \bmod p, \\ acp - bcp(c+1)/2 \equiv e(p-1) - bp(p+1)/2, \bmod p^2. \end{cases}$$

Combining the second and third congruences we get

$$bc^2 \equiv b(p+1)/2, \bmod p.$$

Since b can be neither 0 nor a multiple of p, this leaves

$$c^2 \equiv (p+1)/2, \bmod p,$$

which is in general not true. For example, if $p = 5$, c^2 must be 1 or 4. Therefore, *the groups G and $\bar{G}$ are not simply isomorphic.*

URBANA, ILLINOIS,
FEBRUARY, 1930.